

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:18 ; Search time 5179.29 Seconds  
(without alignments)  
9899.970 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccgaagacc.....tctctttggaatgcctag 1183

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_hg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vt:\*
- 30: em\_hg\_hum:\*
- 31: em\_hg\_inv:\*
- 32: em\_hg\_other:\*
- 33: em\_hg\_mus:\*
- 34: em\_hg\_pln:\*
- 35: em\_hg\_rtd:\*
- 36: em\_hg\_man:\*
- 37: em\_hg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_hgo\_hum:\*
- 40: em\_hgo\_mus:\*
- 41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1183	100.0	1183	6	AX443133	AX443133 Sequence
c	1183	100.0	1183	6	AX443135	AX443135 Sequence
3	1181.4	99.9	2567	6	AX452880	AX452880 Sequence
4	1152.4	97.4	2788	6	AX780857	AX780857 Sequence
5	1144.8	96.8	2538	9	BC042041	BC042041 Homo sapi
6	825.8	69.8	2415	9	AX025645	AX025645 Homo sapi
7	784.4	66.3	786	6	AX511153	AX511153 Sequence
8	784.4	66.3	786	6	AX572845	AX572845 Sequence
9	784.4	66.3	786	9	AF290985	AF290985 Homo sapi
10	784.4	66.3	786	9	AF326353	AF326353 Homo sapi
11	661.8	55.9	737	6	AX511155	AX511155 Sequence
12	661.8	55.9	737	9	AF290986	AF290986 Homo sapi
13	563	47.6	1384	10	BC052655	BC052655 Mus muscu
14	561.4	47.5	1348	6	AX511150	AX511150 Sequence
15	561.4	47.5	1348	10	AF287467	AF287467 Mus muscu
16	559.8	47.3	1321	10	AF434990	AF434990 Mus muscu
17	535.2	45.2	777	6	AX511151	AX511151 Sequence
c	356.2	30.1	66741	9	HS460J8	AL031662 Human DNA
19	356.2	30.1	145833	2	AC026539	AC026539 Homo sapi
20	212.6	18.0	1571	5	BC056035	BC056035 Xenopus l
21	159	13.4	2021	9	HSU44403	U44403 Human Src-1
22	159	13.4	3090	9	HSU44403	HSU44403 Homo sapi
23	158.6	13.4	2109	6	HSU30473	HSU30473 Homo sapi
24	157.4	13.3	1076	9	BC007042	BC007042 Homo sapi
25	157.4	13.3	1870	9	BC007042	BC007042 Homo sapi
26	157.4	13.3	2665	6	AX333017	AX333017 Sequence
27	157.4	13.3	2665	9	D89077	D89077 Homo sapien
c	156.2	13.2	145068	9	HSU977B1	AL050318 Human DNA
29	156.2	13.2	145833	2	AC026539	AC026539 Homo sapi
c	137.6	11.6	849	5	AY278230	AY278230 Gallus Ga
31	132	11.2	841	10	AY079449	AY079449 Mus muscu
32	132	11.2	841	10	AY079450	AY079450 Mus muscu
33	132	11.2	921	10	MMU29056	MMU29056 Mus muscu
34	132	11.2	2648	10	BC032922	BC032922 Mus muscu
35	132	11.2	2773	10	MMU131777	MMU131777 Mus muscu
c	129.8	11.0	299947	2	AC125701	AC125701 Rattus no
37	129.8	11.0	322940	2	AC123560	AC123560 Rattus no
38	128.8	10.9	1631	10	AY217759	AY217759 Rattus no
c	128.8	10.9	163240	10	AL935150	AL935150 Mouse DNA
40	126	10.7	3701	5	CHTKCKL	J03579 Chicken tyr
41	126	10.7	4073	5	GGTKI	X60380 G.gallus tk
42	125.2	10.6	2703	10	MUSLYNB	M57697 Mouse lyn B
43	125.2	10.6	2766	10	MUSLYNA	M57696 Mouse lyn A
44	123.6	10.4	2041	10	RNAF000300	AF000300 Rattus no
45	123.6	10.4	2044	10	MUSLYN	M64608 Mouse lyn p

## ALIGNMENTS

RESULT 1	AX443133	1183 bp	DNA	linear	PAT 02-JUL-2002
AX443133	Sequence 74 from Patent WO0216599.				
LOCUS	AX443133				
DEFINITION	Sequence 74 from Patent WO0216599.				
ACCESSION	AX443133				
VERSION	AX443133.1	GI:21690555			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE					
AUTHORS	Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.				

TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES Location/Qualifiers  
source 1..1183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1183; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 1.2e-305;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTAGAGCTCAAGGACCCACCGCTGTGTCTCTGTGACAGAGCTCAAGAGCCCTGGG 60  
Db 1 AGCTAGAGCTCAAGGACCCACCGCTGTGTCTCTGTGACAGAGCTCAAGAGCCCTGGG 60

QY 61 CTTTCCCTCCCTGGCTGGCTGTGTCTGGAGGGTTCCCGAGTCCAGATCCCTAAGGAG 120  
Db 61 CTTTCCCTCCCTGGCTGGCTGTGTCTGGAGGGTTCCCGAGTCCAGATCCCTAAGGAG 120

QY 121 CATGGGCGAGCTGATCCATCCCTGGTGTACAACTGTGATGTCAGACAGATGCTGAGCT 180  
Db 121 CATGGGCGAGCTGATCCATCCCTGGTGTACAACTGTGATGTCAGACAGATGCTGAGCT 180

QY 181 ACCCAAAACCAACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAGTTCTGGGTG 240  
Db 181 ACCCAAAACCAACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAGTTCTGGGTG 240

QY 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTCCA 300  
Db 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTCCA 300

QY 301 GCAGAGCATGCTGTCCAGCAGCTGTCTCCAGGCTTTGATGACAAACCAATTTC 360  
Db 301 GCAGAGCATGCTGTCCAGCAGCTGTCTCCAGGCTTTGATGACAAACCAATTTC 360

QY 361 CTCGATGATGCTGTCTGAGTCTCTGCTGAGGAAATGGAAGTCTGCCAGCAGAG 420  
Db 361 CTCGATGATGCTGTCTGAGTCTCTGCTGAGGAAATGGAAGTCTGCCAGCAGAG 420

QY 421 AAAATCTCTGCAAGCCAGCTTGATGCTCTCTGCAAGCCAGGACCTGTGACCAT 480  
Db 421 AAAATCTCTGCAAGCCAGCTTGATGCTCTCTGCAAGCCAGGACCTGTGACCAT 480

QY 481 GGAAGCAGAGAGAGCAAGCCACAGCCGTCGCTGGCCCTTCCCGCAGGTGGGCC 540  
Db 481 GGAAGCAGAGAGAGCAAGCCACAGCCGTCGCTGGCCCTTCCCGCAGGTGGGCC 540

QY 541 GGGCGAGCTGTGCTGAGACTCGGGAGCCATGACCATGCTCTGAGGATGGAGACTG 600  
Db 541 GGGCGAGCTGTGCTGAGACTCGGGAGCCATGACCATGCTCTGAGGATGGAGACTG 600

QY 601 GTGGACGCTGTCTGTAAGTCTCAGGACAGAGATATACATCCCGCAGCTCCAGTGGG 660  
Db 601 GTGGACGCTGTCTGTAAGTCTCAGGACAGAGATATACATCCCGCAGCTCCAGTGGG 660

QY 661 CAAAGTCTCCATGGGTGCTGTATGAGGCTGTGAGGGGAAAGCAGAGGAACTGCT 720  
Db 661 CAAAGTCTCCATGGGTGCTGTATGAGGCTGTGAGGGGAAAGCAGAGGAACTGCT 720

QY 721 GTTGTACTCTGGAAACCTGGAGGGCCCTTCTCATCCGGAGGACGACAGAGAGG 780  
Db 721 GTTGTACTCTGGAAACCTGGAGGGCCCTTCTCATCCGGAGGACGACAGAGAGG 780

QY 781 CTCCTTACTCTGTAGTCCGCTCAGCCGCTTGCATCTCGGACCGGATCAGACACTA 840  
Db 781 CTCCTTACTCTGTAGTCCGCTCAGCCGCTTGCATCTCGGACCGGATCAGACACTA 840

QY 841 CAGGATCCATGCTGCAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTC 900  
Db 841 CAGGATCCATGCTGCAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTC 900

QY 901 ACTCCAGGCGCTGGTGAACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCAA 960  
Db 901 ACTCCAGGCGCTGGTGAACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCAA 960

QY 961 GGAGCCCTGTGTCTCTGACAGAGGCTGGCCGCTCCCTGCAAGGATATACCCCTACTCT 1020  
Db 961 GGAGCCCTGTGTCTCTGACAGAGGCTGGCCGCTCCCTGCAAGGATATACCCCTACTCT 1020

QY 1021 GACTGTGACAGAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTTCTGA 1080  
Db 1021 GACTGTGACAGAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTTCTGA 1080

QY 1081 AGCTGACACAGGGAGAGTCTCTTCTCASTAGGCTTCCGGAGTCCCTGAGTCTCTA 1140  
Db 1081 AGCTGACACAGGGAGAGTCTCTTCTCASTAGGCTTCCGGAGTCCCTGAGTCTCTA 1140

QY 1141 CATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTAG 1183  
Db 1141 CATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTAG 1183

RESULT 2  
AX443135/c  
LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002  
DEFINITION Sequence 76 from Patent WO0216599.  
ACCESSION AX443135  
VERSION AX443135.1 GI:21690556  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Burgess, C.E., Conley, P.B., Grosse, W.W., Hart, M., Kekuda, R., Shimkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E., Topper, J.N. and Yang, R.B.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES Location/Qualifiers  
source 1..1183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1183; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 1.2e-305;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTAGAGCTCAAGGACCCACCGCTGTGTCTCTGTGACAGAGCTCAAGAGCCCTGGG 60  
Db 1183 AGCTAGAGCTCAAGGACCCACCGCTGTGTCTCTGTGACAGAGCTCAAGAGCCCTGGG 1124

QY 61 CTTTCCCTCCCTGGCTGGCTGTGTCTGGAGGGTTCCCGAGTCCAGATCCCTAAGGAG 120  
Db 1123 CTTTCCCTCCCTGGCTGGCTGTGTCTGGAGGGTTCCCGAGTCCAGATCCCTAAGGAG 1064

QY 121 CATGGGCGAGCTGATCCATCCCTGGTGTACAACTGTGCTGACGACAGATGCTGAGCT 180  
Db 1063 CATGGGCGAGCTGATCCATCCCTGGTGTACAACTGTGCTGACGACAGATGCTGAGCT 1004

QY 181 ACCCAAAACCAACCTAGCTCTCCCTGAAGTCTCCAGGCTGAGAGGTTCTGGGTG 240  
Db 1003 ACCCAAAACCAACCTAGCTCTCCCTGAAGTCTCCAGGCTGAGAGGTTCTGGGTG 944

QY 241 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTCCA 300  
Db 943 TCTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTCCA 884

QY 301 GCAGAGCATGCTGTGACAGAGCTGTCTTCCAGGCTTTGATGACAAACCAATTTC 360



858 CAGGATCCACTGCTTGGACAAATGGCTGGCTGTATCATCTCACCGCGCTCACCTTCCCTC 917  
 901 ACTCCAGGCCCTGGTGGACCACTACTCTGAGCTGGGGATGACATCTGCTGCTACTCAA 960  
 918 ACTCCAGGCCCTGGTGGACCACTACTCTGAGCTGGGGATGACATCTGCTGCTACTCAA 977  
 961 GGAGCCCTGTGCTCTGTCAGAGGGCTGGCCGCTCCCTGGCAAGATATACCCCTACTCTG 1020  
 978 GGAGCCCTGTGCTCTGTCAGAGGGCTGGCCGCTCCCTGGCAAGATATACCCCTACTCTG 1037  
 1021 GACTGTCCAGAGGACACCACTCACTCACTGGAAGAGCTGGAGCTCCCTCTGTTTCTGA 1080  
 1038 GACTGTCCAGAGGACACCACTCACTCACTGGAAGAGCTGGAGCTCCCTCTGTTTCTGA 1097  
 1081 AGCTGCCACAGGGAGGAGTCTCTTCTCAGTGGGCTCTCCGGAGTCCCTCAGCTTCTA 1140  
 1098 AGCTGCCACAGGGAGGAGTCTCTTCTCAGTGGGCTCTCCGGAGTCCCTCAGCTTCTA 1157  
 1141 CATCAGCCCTGAATCAGAGGCTGTCTCTTTGGATGATGCCTAG 1193  
 1158 CATCAGCCCTGAATCAGAGGCTGTCTCTTTGGATGATGCCTAG 1200

RESULT 4  
 AX780857 2788 bp DNA linear PAT 14-JUL-2003  
 LOCUS Sequence 3014 from Patent WO03039443.  
 DEFINITION AX780857  
 ACCESSION AX780857  
 VERSION AX780857.1 GI:32697851  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
 Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.  
 Novel genetic markers for leukemias  
 Patent: WO 03039443-A 3014 15-MAY-2003;  
 Deutsches Krebsforschungszentrum (DE); Haferlach, Torsten,  
 Ludwig-Maximilians-Universitaet Muenchen (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)  
 PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)  
 Location/Qualifiers  
 1..2788  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 97.4%; Score 1152.4; DB 6; Length 2788;  
 Best Local Similarity 98.4%; Pred. No. 1.8e-297;  
 Matches 1153; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
 12 CAAGGACCCAGCGCTGTCTCTGTGACAGAGTCAAGAGGCGCTGGGCTTCCCTCCC 71  
 1 CAAGGACCCAGCGCTGTCTCTGTGACAGAGTCAAGAGGCGCTGGGCTTCCCTCCC 60  
 72 TGCTCGGCTGTCTTGGAGGGTTCCCGAGTCCAGATCCCTAAGAGCATGGGGCAGC 131  
 61 TGCTCGGCTGTCTTGGAGGGTTCCCGAGTCCAGATCCCTAAGAGCATGGGGCAGC 120  
 132 TGATCCATCCCTGGTGTACAACTGCTGCTGACACAGAGTCTGAGCTACCCAAACCA 191  
 121 TGATCCATCCCTGGTGTACAACTGCTGCTGACACAGAGTCTGAGCTACCCAAACCA 180  
 192 CACCTAGCCTCTCCCTGAAGATCTCTCCAGGCTGAGAGATTTCTGGGTGTCTAGGACA 251  
 181 CACCTAGCCTCTCCCTGAAGATCTCTCCAGGCTGAGAGATTTCTGGGTGTCTAGGACA 240  
 252 AGGACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTTCCAGCAGAGCATG 311  
 241 AGGACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTTCCAGCAGAGCATG 300

312 CGTCTCAGCAGAGTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCGATGATGT 371  
 301 CGTCTCAGCAGAGTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCGATGATGT 360  
 372 GCTTCTGAGTCTCTGCTGAGGAAATGGGAAGTCTGCCAGCAGAGAAAATCTCTGC 431  
 361 GCTTCTGAGTCTCTGCTGAGGAAATGGGAAGTCTGCCAGCAGAGAAAATCTCTGC 420  
 432 CAAGCCCAAGCTTCTGCTGAGGAAATGGGAAGTCTGCCAGCAGAGAAAATCTCTGC 491  
 421 CAAGCCCAAGCTTCTGCTGAGGAAATGGGAAGTCTGCCAGCAGAGAAAATCTCTGC 480  
 492 GAAGCAAGCCACAGCGCTGGCGAGTTTCCCGGAGAGTGGCCCGCCGAGCTGT 551  
 481 GAAGCAAGCCACAGCGCTGGCGAGTTTCCCGGAGAGTGGCCCGCCGAGCTGT 540  
 552 CGCTCAGACTCGGGAGCCATGACCATCTCTCTGAGGATGGAGACTGTGGACGGTGC 611  
 541 CGCTCAGACTCGGGAGCCATGACCATCTCTCTGAGGATGGAGACTGTGGACGGTGC 600  
 612 TGCTTGAAGTCTCAGGACAGAGTATACATCCCGAGGCTCCAGTGGGCAAGTCTCCC 671  
 601 TGCTTGAAGTCTCAGGACAGAGTATACATCCCGAGGCTCCAGTGGGCAAGTCTCCC 660  
 672 ATGGGTGGCTGTATGAGGCGCTGAGCAGGAGAAAGAGAGAACTGTGTGTACTCTG 731  
 661 ATGGGTGGCTGTATGAGGCGCTGAGCAGGAGAAAGAGAGAACTGTGTGTACTCTG 720  
 732 GGAACCCCTGGAGGGCGCTTCTCTATCCCGGAGAGCCAGACAGGAGAGGCTTCTTACTCTC 791  
 721 GGAACCCCTGGAGGGCGCTTCTCTATCCCGGAGAGCCAGACAGGAGAGGCTTCTTACTCTC 780  
 792 TGTCAGTCCGCTCAGCGCGCTGATCTCTGGAGCCGATCAGACACTACAGGATCCACT 851  
 781 TGTCAGTCCGCTCAGCGCGCTGATCTCTGGAGCCGATCAGACACTACAGGATCCACT 840  
 852 GCTTGAACATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCCTCAGGCGCC 911  
 841 GCTTGAACATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCCTCAGGCGCC 900  
 912 TGTGTGACCATTAATCTGAGCTGGCGGATGACATCTGTGTGCTTCTCAAGAGAGCCCTGTG 971  
 901 TGTGTGACCATTAATCTGAGCTGGCGGATGACATCTGTGTGCTTCTCAAGAGAGCCCTGTG 960  
 972 TCTTGCAGAGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGA 1031  
 961 TCTTGCAGAGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGA 1020  
 1032 GGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTTCTGAAGCTGGCACAG 1091  
 1021 GGACACCACTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTTCTGAAGCTGGCACAG 1080  
 1092 GGAGAGGCTCTTCTCAGTGGAGGCTTCCGGAGTCCCTCAGCTTCTACATCAGCCTGA 1151  
 1081 GGAGAGGCTCTTCTCAGTGGAGGCTTCCGGAGTCCCTCAGCTTCTACATCAGCCTGA 1140  
 1152 ATCAGCAGGCTGTCTCTTTGGATGATGCCTAG 1183  
 1141 ATCAGCAGGCTGTCTCTTTGGATGATGCCTAG 1172

RESULT 5  
 BC042041 2538 bp mRNA linear PRI 07-OCT-2003  
 LOCUS Homo sapiens Szc-like-adaptor 2, transcript variant 1, mRNA (cdna  
 DEFINITION clone MGC:49845 IMAGE:4429896), complete cds.  
 ACCSSION BC042041  
 VERSION BC042041.1 GI:27469842  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```
Db CATCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGAACATGGCTGCTGTACA 840
QY TCTCACCGCCCTCACCTCCCTCACTCCAGGCCCTGGTGACCACTTACTCTGAGCTGG 935
Db TCTCACCGCCCTCACCTCCCTCACTCCAGGCCCTGGTGACCACTTACTCTGAGCTGG 900
QY CGGATGACATCTGCTGCTACTCAAGAGAGCCCTGTGCTTGCAGAGAGGCTGGCCCGCTCC 995
Db CGGATGACATCTGCTGCTACTCAAGAGAGCCCTGTGCTTGCAGAGAGGCTGGCCCGCTCC 960
QY CTGCAAGGATATACCCCTACCTGCTGACTGTGAGAGACACCACTCAACTGGAAGAGC 1055
Db CTGCAAGGATATACCCCTACCTGCTGACTGTGAGAGACACCACTCAACTGGAAGAGC 1020
QY TGGACAGCTCCCTCTGCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGG 1115
Db TGGACAGCTCCCTCTGCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGG 1080
QY GTCTCCGGGAGTCCCTGAGCTTCTACATCAGCTGAATGAGGAGGCTGCTCTTTGGATG 1175
Db GTCTCCGGGAGTCCCTGAGCTTCTACATCAGCTGAATGAGGAGGCTGCTCTTTGGATG 1140
QY 1176 ATGCTAG 1183
Db 1141 ATGCTAG 1148

RESULT 6
AK025645 2415 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA: FLJ21992 fis, clone HEP06554.
DEFINITION AK025645
ACCESSION AK025645
VERSION AK025645.1 GI:10438227
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,A., Oka,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEPO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2415)
AUTHORS Sugano,S., Suzuki,Y., Oka,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEPO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES Location/Qualifiers
source 1..2415
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
clone="HEP06554"
cell_line="HepG2"
cell_type="hepatoma"
clone_lib="HEP"
note="cloning vector pME18SFL3"
61..846
note="unnamed protein product"

CDS
```

```
/codon_start=1
/protein_id="BAB15201.1"
/db_xref="GI:10438228"
/translation="MGSLPGRKSLPSPSLSSVQGGQPVMTAEERSKATAVALGSPF
AGGPAELSLRLGPELTIIVSDGDWMTVLSEVSGREYNI PSVHAKVSHGWLVEGLSRE
KAEILLLLPNPGGAFLIRSDTERGSYLSVRLSRPASWDRI RHVRIHCLDNGWLVI
SPRLTFPSLQALVDHYSELADIDICLLKPCVLQRAGPLKGDJPLPVTQRTLNWK
ELDSSLTFSEAAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"

ORIGIN
Query Match 69.8%; Score 825.8; DB 9; Length 2415;
Best Local Similarity 99.8%; Pred. No. 5.4e-210;
Matches 827; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 355 ATTCCCTCGATGATGCTCTTCTGAGTCTCTCTGAGGAACAATGGAAGTCTGCCAG 414
Db 18 ATTCCCTCGATGATGCTCTTCTGAGTCTCTCTGAGGAACAATGGAAGTCTGCCAG 77
QY 415 CAGAAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTCTCCAAGCCAGGACCTGT 474
Db 78 CAGAAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTCTCCAAGCCAGGACCTGT 137
QY 475 GACCATGGAAGCAGAGAGAGCAAGCCACAGCCCTGSCCTGGCAGTTCCTCCGGCAGG 534
Db 138 GACCATGGAAGCAGAGAGAGCAAGCCACAGCCCTGSCCTGGCAGTTCCTCCGGCAGG 197
QY 535 TGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGG 594
Db 198 TGGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGG 257
QY 595 AGACTGTGGAGCGGTGCTGTCTGAAGTCTCAGGCGAGAGAGTATAACATCCCGAGGCTCA 654
Db 258 AGACTGTGGAGCGGTGCTGTCTGAAGTCTCAGGCGAGAGAGTATAACATCCCGAGGCTCA 317
QY 655 CGTGGGCAAGCTCCCATGGGTGCTGTATGAGGCGCTGAGCAGGAGAGAGAGAGGA 714
Db 318 CTTGGCCAAAGTCTCCATGGGTGCTGTATGAGGCGCTGAGCAGGAGAGAGAGAGGA 377
QY 715 ACTGCTGTTTACCTGGGAAACCCCTGGAGGGCCCTTCCTCATCCGGGAGAGCCAGACAG 774
Db 378 ACTGCTGTTTACCTGGGAAACCCCTGGAGGGCCCTTCCTCATCCGGGAGAGCCAGACAG 437
QY 775 GAGAGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCTGTCATCTGGACCGGATCAG 834
Db 438 GAGAGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCTGTCATCTGGACCGGATCAG 497
QY 835 AACTACAGGATCCACTGCTTTGACAATGGTGGCTGTACATCTCACCGCGCTCACCTTT 894
Db 498 AACTACAGGATCCACTGCTTTGACAATGGTGGCTGTACATCTCACCGCGCTCACCTTT 557
QY 895 CCCTTCACTCCAGGCCCTGGTGAGCACTTACTCTGAGCTGGCGGATGACATCTGCTGCT 954
Db 558 CCCTTCACTCCAGGCCCTGGTGAGCACTTACTCTGAGCTGGCGGATGACATCTGCTGCT 617
QY 955 ACTCAAGAGAGCCCTGTGCTCTGACAGAGGCTGGCCCGCTCCCTGGCAAGATATACCCCT 1014
Db 618 ACTCAAGAGAGCCCTGTGCTCTGACAGAGGCTGGCCCGCTCCCTGGCAAGATATACCCCT 677
QY 1015 ACTGTGACTGTGCAGAGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTT 1074
Db 678 ACCTGTGACTGTGCAGAGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTT 737
QY 1075 TTCTGAAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAG 1134
Db 738 TTCTGAAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAG 797
QY 1135 CTCTTACATCAGGCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183
Db 798 CTCTTACATCAGGCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 846

RESULT 7
AXS11153
```

LOCUS AX511153 786 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 4 from Patent WO0242452.  
ACCESSION AX511153  
VERSION AX511153.1 GI:23392046  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Mcglade,J.C. and Loreto,M.P.  
TITLE Adapter gene  
JOURNAL Patent: WO 0242452-A 4 30-MAY-2002;  
The Hospital for Sick Children (CA)  
FEATURES  
source  
1..786  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 66.3%; Score 784.4; DB 6; Length 786;  
Best Local Similarity 99.9%; Pred. No. 6.9e-199;  
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 457  
DB 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60  
QY 458 CAAGCCAGGACCTGTGACCATGAGCAGAGAGAAAGGCCACAGCGTGGCCCTG 517  
DB 61 CAAGCCAGGACCTGTGACCATGAGCAGAGAGAAAGGCCACAGCGTGGCCCTG 120  
QY 518 GGCAGTTTCCCGCAGGTCGCCCGCCGAGCTGCGTGGAGCTCGGGAGCCATTGACC 577  
DB 121 GGCAGTTTCCCGCAGGTCGCCCGCCGAGCTGCGTGGAGCTCGGGAGCCATTGACC 180  
QY 578 ATCTCTCTGAGGATGAGAGCTGTGGACGGTGTCTGTGAAGTCTCAGGCAGAGATAT 637  
DB 181 ATCTCTCTGAGGATGAGAGCTGTGGACGGTGTCTGTGAAGTCTCAGGCAGAGATAT 240  
QY 638 AACATCCCAGGTCACGTCGGCAGAGTCTCCCATGGTGGTGTATGAGGCGCTGAGC 697  
DB 241 AACATCCCAGGTCACGTCGGCAGAGTCTCCCATGGTGGTGTATGAGGCGCTGAGC 300  
QY 698 AGGAGAAAGCAGAGAACTGCTGTGTTTACTTGGAACTTGGAGGCGCTTCTCTCATC 757  
DB 301 AGGAGAAAGCAGAGAACTGCTGTGTTTACTTGGAACTTGGAGGCGCTTCTCTCATC 360  
QY 758 CGGAGAGCCAGACAGAGAGGCTTCTACTCTGTGTCAGTCCGCTCAGCGCCCTGCA 817  
DB 361 CGGAGAGCCAGACAGAGAGGCTTCTACTCTGTGTCAGTCCGCTCAGCGCCCTGCA 420  
QY 818 TCCTGGGACCGATCAGACATACAGGATCCACTGCTTGAATGCTGCTGTATCATC 877  
DB 421 TCCTGGGACCGATCAGACATACAGGATCCACTGCTTGAATGCTGCTGTATCATC 480  
QY 878 TCACCGCCCTCACTTCCCTCACTCAGGCGCTGTGGACCACTTACTTGAGCTGGCG 937  
DB 481 TCACCGCCCTCACTTCCCTCACTCAGGCGCTGTGGACCACTTACTTGAGCTGGCG 540  
QY 938 GATGACATCTGCTGCTACTCAGAGGCGCTGTGCTGTCAGAGGCTGCGCGCTCCCT 997  
DB 541 GATGACATCTGCTGCTACTCAGAGGCGCTGTGCTGTCAGAGGCTGCGCGCTCCCT 600  
QY 998 GGCAGGATATACCCCTACCTGTGATGTGAGAGGACACCACTCAACTGGAAGAGCTG 1057  
DB 601 GGCAGGATATACCCCTACCTGTGATGTGAGAGGACACCACTCAACTGGAAGAGCTG 660  
QY 1058 GACAGCTCCCTCTGTTTCTGAACTGCGCAGAGGAGGATCTCTCTCAGTGAGGCT 1117  
DB 661 GACAGCTCCCTCTGTTTCTGAACTGCGCAGAGGAGGATCTCTCTCAGTGAGGCT 720

QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGGATGAT 1177  
DB 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTTTGGATGAT 780  
QY 1178 GCCTAG 1183  
DB 781 GCCTAG 786  
RESULT 8  
AX512845  
LOCUS AX512845 786 bp DNA linear PAT 29-NOV-2002  
DEFINITION Sequence 1 from Patent WO02055707.  
ACCESSION AX512845  
VERSION AX512845.1 GI:26004935  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Holland,S.J., Mendenhall,M.K., Pardo,J., Spencer,C., Fu,A.C.,  
Luo,Y., Payan,D.G., Mancebo,H.S., Wu,J., Zhou,X., Shen,M.,  
Liao,X.C. and Sheng,N.  
TITLE Cloning of an inhibitor of antigen-receptor signaling by a  
retroviral-based functional screen  
JOURNAL Patent: WO 02055707-A 1 18-JUL-2002;  
Rigel Pharmaceuticals, Inc. (US)  
FEATURES  
source  
1..786  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..786  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD58542.1"  
/db\_xref="GI:26004936"  
/db\_xref="RMTREMBL:CAD58542"  
/translation="MGSLPSRKSLSPSLSSVSGQGVMTMEERSKATAVALGSFP  
AGGAELESLRLEPIITVSEDDQWTVLSEVSGREYNIPSHVAKVSHGWLVEGLSRE  
KABELLLPQNPFGAFLIRSEQTRGYSLSVLRSPASWDRIHYRHLCDNGWLYI  
SPRLTFQLADHYLSLADDDICCLKEPCVLQAGPLFGKDIPLPVTQVQPLNWK  
ELDSSLFLSEATGEBESLLSEGLRESLSFYISLNDEAVSLDDA"  
ORIGIN  
Query Match 66.3%; Score 784.4; DB 6; Length 786;  
Best Local Similarity 99.9%; Pred. No. 6.9e-199;  
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 457  
DB 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60  
QY 458 CAAGCCAGGACCTGTGACCATGAGCAGAGAGAAAGGCCACAGCGTGGCCCTG 517  
DB 61 CAAGCCAGGACCTGTGACCATGAGCAGAGAGAAAGGCCACAGCGTGGCCCTG 120  
QY 518 GGCAGTTTCCCGCAGGTCGCCCGCCGAGCTGCGTGGAGCTCGGGAGCCATTGACC 577  
DB 121 GGCAGTTTCCCGCAGGTCGCCCGCCGAGCTGCGTGGAGCTCGGGAGCCATTGACC 180  
QY 578 ATCTCTCTGAGGATGAGAGCTGTGGACGGTGTCTGTGAAGTCTCAGGCAGAGATAT 637  
DB 181 ATCTCTCTGAGGATGAGAGCTGTGGACGGTGTCTGTGAAGTCTCAGGCAGAGATAT 240  
QY 638 AACATCCCAGGTCACGTCGGCAGAGTCTCCCATGGTGGTGTATGAGGCGCTGAGC 697  
DB 241 AACATCCCAGGTCACGTCGGCAGAGTCTCCCATGGTGGTGTATGAGGCGCTGAGC 300  
QY 698 AGGAGAAAGCAGAGAACTGCTGTGTTTACTTGGAACTTGGAGGCGCTTCTCTCATC 757  
DB 301 AGGAGAAAGCAGAGAACTGCTGTGTTTACTTGGAACTTGGAGGCGCTTCTCTCATC 360

```
QY 758 CGGAGAGCCAGACACAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCTGCA 817
Db 361 CGGAGAGCCAGACACAGAGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCTGCA 420
QY 818 TCCTGGACCGGATCAGACACTACAGATCCACTGCTTGCCTTGACATGGCTGTACATC 877
Db 421 TCCTGGACCGGATCAGACACTACAGATCCACTGCTTGCCTTGACATGGCTGTACATC 480
QY 878 TCACCGGCTCTACCTTCCCTCACTCCAGGCTTGTGTGACCACTTACTCTGAGCTGGG 937
Db 481 TCACCGGCTCTACCTTCCCTCACTCCAGGCTTGTGTGACCACTTACTCTGAGCTGGG 540
QY 938 GATGACATCTGCTCTACTCAGGAGCCCTGTCTCTGACAGAGGCTGCCCCCTCCCT 997
Db 541 GATGACATCTGCTCTACTCAGGAGCCCTGTCTCTGACAGAGGCTGCCCCCTCCCT 600
QY 998 GGCAGGATATACCCCTACCTGTGACGTGTGACAGGACACCACTCAACTGGAAGAGCTG 1057
Db 601 GGCAGGATATACCCCTACCTGTGACGTGTGACAGGACACCACTCAACTGGAAGAGCTG 660
QY 1058 GACAGCTCCCTCTCTCTTCTGAACTGCGACAGGAGGAGTCTCTTCTCAGTGAGGT 1117
Db 661 GACAGCTCCCTCTCTCTTCTGAACTGCGACAGGAGGAGTCTCTTCTCAGTGAGGT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGACAGGCTGTCTCTTTGGATGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGACAGGCTGTCTCTTTGGATGAT 780
QY 1178 GCCTAG 1183
Db 781 GCCTAG 786

RESULT 9
AF290985 786 bp mRNA linear PRI 21-JAN-2003
LOCUS Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
DEFINITION AF290985
ACCESSION AF290985.1 GI:17351920
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Cloning and characterization of human Src-like adaptor protein 2
and a novel splice isoform, SLAP-2-v
JOURNAL Oncogene 22 (2), 266-273 (2003)
MEDLINE 22415750
PubMed 12527895
REFERENCE 2 (bases 1 to 786)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1..786-
/note="SLAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAL38197.1"
/db_xref="GI:17351921"
/translation="WGSLPGRKSLPSPSLSSVQGGPVTMEARSKATAVALGSFP
AGGPAELSLRIGELTIVSEDCDWTVLSEVSGREYNIPSHVAKVSHGWLIEGLSRE
KAEELLULPFGNFGAFLIREQTRGRGSYSVLRUSPASMWRIRHYNHICLDNGWLYI
```

```
SPRLTFPSLQALVDHYSELADDDICLLKPCVLTQAGPLPKGDIPLPVTQVTRPLNWK
ELDSSLFSEAAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"

ORIGIN
Query Match 66.3%; Score 784.4; DB 9; Length 786;
Best Local Similarity 99.9%; Pred. No. 6.9e-199;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 398 ATGGAACTTGCCTCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCCTCTGTC 457
Db 1 ATGGAACTTGCCTCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCCTCTGTC 60
QY 458 CAGGCCAGGACCTGTGACCATGAGACAGAGAGAGGACACAGCCGTGSCCCTG 517
Db 61 CAGGCCAGGACCTGTGACCATGAGACAGAGAGAGGACACAGCCGTGSCCCTG 120
QY 518 GCGATTTCCCGCGAGGTGCGCCGCGAGCTGTCGCTGAGACTCGGGAGACCAATTGACC 577
Db 121 GCGATTTCCCGCGAGGTGCGCCGCGAGCTGTCGCTGAGACTCGGGAGACCAATTGACC 180
QY 578 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGAAGTCTCAGGCAGAGAGTAT 637
Db 181 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGAAGTCTCAGGCAGAGAGTAT 240
QY 638 AACATCCCCAGCGTCCACGTCGGCAAGTCTCCCATGGGTGCTGTATGAGGGCTTGAGC 697
Db 241 AACATCCCCAGCGTCCACGTCGGCAAGTCTCCCATGGGTGCTGTATGAGGGCTTGAGC 300
QY 698 AGGAGAAAGCAGAGAGAACTGTTTACTGGGAACCTCGAGGGGCTTCTCTCATC 757
Db 301 AGGAGAAAGCAGAGAGAACTGTTTACTGGGAACCTCGAGGGGCTTCTCTCATC 360
QY 758 CGGAGAGCCAGACCCAGGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCTGCA 817
Db 361 CGGAGAGCCAGACCCAGGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCTGCA 420
QY 818 TCCTGGACCGGATCAGACACTACAGATCCACTGCTTGCCTTGACATGGCTGCTACATC 877
Db 421 TCCTGGACCGGATCAGACACTACAGATCCACTGCTTGCCTTGACATGGCTGCTACATC 480
QY 878 TCACCGGCTCTACCTTCCCTCACTCCAGGCTTGTGTGAGCCATTACTCTGAGTGGCG 937
Db 481 TCACCGGCTCTACCTTCCCTCACTCCAGGCTTGTGTGAGCCATTACTCTGAGTGGCG 540
QY 938 GATGACATCTGCTCTACTCAGAGGCTGTGTCTGACAGAGGCTGCCCCCTCCCT 997
Db 541 GATGACATCTGCTCTACTCAGAGGCTGTGTCTGACAGAGGCTGCCCCCTCCCT 600
QY 998 GGCAGGATATACCCCTACCTGTGACGTGTGACAGGACACCACTCAACTGGAAGAGCTG 1057
Db 601 GGCAGGATATACCCCTACCTGTGACGTGTGACAGGACACCACTCAACTGGAAGAGCTG 660
QY 1058 GACAGCTCCCTCTCTTCTTCTGAACTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGT 1117
Db 661 GACAGCTCCCTCTCTTCTTCTGAACTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGACAGGCTGTCTCTTTGGATGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGACAGGCTGTCTCTTTGGATGAT 780
QY 1178 GCCTAG 1183
Db 781 GCCTAG 786

RESULT 10
AF326353 786 bp mRNA linear PRI 08-NOV-2001
LOCUS Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
DEFINITION AF326353
ACCESSION AF326353
VERSION AF326353.1 GI:16797891
KEYWORDS Homo sapiens (human)
SOURCE
```

Db	421	TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGA	CAATGGCTGGCTGTATCATC	480
QY	878	TCACGCGGCTCACCTTCCCTTCACTCCAGGCGCTGGTGGACCACTTACTGTAGCTGGCG	937	
Db	481	TCACGCGGCTCACCTTCCCTTCACTCCAGGCGCTGGTGGACCACTTACTGTAGCTGGCG	540	
QY	938	GATGACATCTGCTGCCTACTCAAGGAGCCCTGTGCTCTGCAGAGGGCTGGCCCGCTCCCT	997	
Db	541	GATGACATCTGCTGCCTACTCAAGGAGCCCTGTGCTCTGCAGAGGGCTGGCCCGCTCCCT	600	
QY	998	GGCAAGGATATACCCCTACTCTGTGACTGTGCAGAGACACCACTCAACTGGAAGAGCTG	1057	
Db	601	GGCAAGGATATACCCCTACTCTGTGACTGTGCAGAGACACCACTCAACTGGAAGAGCTG	660	
QY	1058	GACAGCTCCCTCTGTTTCTGAAGTGGCCACAGGGAGGAGTCTTCTCACTGAGGGT	1117	
Db	661	GACAGCTCCCTCTGTTTCTGAAGTGGCCACAGGGAGGAGTCTTCTCACTGAGGGT	720	
QY	1118	CTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGAT	1177	
Db	721	CTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGAT	780	
QY	1178	GCCTAG 1183		
Db	781	GCCTAG 786		
RESULT 11				
AXS11155	AXS11155	Sequence 6 from Patent WO0242452.		
LOCUS	AXS11155	GI:23392047		
DEFINITION	AXS11155			
ACCESSION	AXS11155			
VERSION	AXS11155.1			
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	McGladre, J.C. and Loreto, M.P.			
JOURNAL	Patent: WO 0242452-A 6 30-MAY-2002;			
FEATURES	The Hospital for Sick Children (CA)			
source	Location/Qualifiers			
	1. .737			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	55.9%;	Score 661.8;	DB 6;	Length 737;
Best Local Similarity	93.3%;	Pred. No. 4,7e-166;		
Matches 734;	Conservative	0;	Mismatches 2;	Indels 51; Gaps 2;
QY	398	ATGGGAAGTCTGCCAGCAGAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	457	
Db	1	ATGGGAAGTCTGCCAGCAGAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC	60	
QY	458	CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCGCTGGCCCTG	517	
Db	61	CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCGCTGGCCCTG	120	
QY	518	GGCAGTTTCCGCGCAGTGGCCCGCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC	577	
Db	121	GGCAGTTTCCGCGCAGTGGCCCGCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC	180	
QY	578	ATCGTCTCTGAGATGGAGACTGGTGGACGGTCTGTCTGAAGTCTCAGCAGAGATAT	637	
Db	181	ATCGTCTCTGAGATGGAGACTGGTGGACGGTCTGTCTGAAGTCTCAGCAGAGATAT	240	
QY	638	AACATCCGAGCTCAGCTGGGCAAGTCTCCATGGTGGCTGTATGAGGCGCTGAC	697	

Db 241 AACATCCCGAGCGTCCACGTCGGCAAGTCTCCATGGGTGGTGTATGAGGCGCTGAGC 300

Qy 698 AGGAGAGAAACAGAGGAACCTGCTGTGTTTACCTGGGAACCCCTGGAGGGCGCTTCTCTCATC 757

Db 301 AGGAGAGAAACAGAGGAACCTGCTGTGTTTACCTGGGAACCCCTGGAGGGCGCTTCTCTCATC 360

Qy 758 GGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTCTGTCAGTCCGCTCAGCCGCTGCA 817

Db 361 CGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTCTGTCAGTCCGCTCAGCCGCTGCA 420

Qy 818 TCCTGGACCGGATCAGACACTCAGGATCCACTGCTTGCACCAATGGCTGGCTGTACATC 877

Db 421 TCCTGGACCGGATCAGACACTCAGGATCCACTGCTTGCACCAATGGCTGTACATC 480

Qy 878 TCACCGGCTCAGCTCCCTCCTACTCCAGGCGCTGTGAGCACTTACTCTGAGCTGGG 937

Db 481 TCACCGGCTCAGCTCCCTCCTACTCCAGGCGCTGTGAGCACTTACTCTT 531

Qy 938 GATGACATCTGCTCCTACTCAAGGAGCCCTGTGTCTCTGTCAGAGGGCTGGCCGCTCCCT 997

Db 532 -----GAGGCTGGCCGCTCCCT 550

Qy 998 GCGAAGATATACCCCTACCTGTGTCAGAGGACACACTCACTCGAAGAGCTG 1057

Db 551 GCGAAGATATACCCCTACCTGTGTCAGAGGACACACTCACTCACTGGAAGAGCTG 610

Qy 1058 GACAGCTCCCTCCGTTTCTGAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAAGGT 1117

Db 611 GACAGCTCCCTCCGTTTCTGAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAAGGT 670

Qy 1118 CTCGGAGTCCCTCAGCTTCTATCAGCTGGAATGA-CGAGGCTGCTCTTTGGATGA 1176

Db 671 CTCGGAGTCCCTCAGCTTCTATCAGCTGGAATGA-CGAGGCTGCTCTTTGGATGA 730

Qy 1177 TGCCTAG 1183

Db 731 TGCCTAG 737

RESULT 12

LOCUS AF290986 737 bp mRNA linear PRI 21-JAN-2003

DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA, complete cds; alternatively spliced.

ACCESSION AF290986

VERSION AF290986.1 GI:17351922

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 737)

AUTHORS Loreto,M.P. and McGlade,C.J.

TITLE Cloning and characterization of human Src-like adaptor protein 2 and a novel splice isoform, SLAP-2-v

JOURNAL Oncogene 22 (2), 266-273 (2003)

MEDLINE 22415750

PUBMED 12527895

REFERENCE 2 (bases 1 to 737)

AUTHORS Loreto,M.P. and McGlade,C.J.

TITLE Direct Submission

JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada

FEATURES

source

1..737

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="20"

/tissue\_type="thymus"

1..533

/note="SLAP-2-v"

/codon\_start=1

CDS

/product="Src-like adaptor protein-2 splice isoform"

/protein\_id="AAL38198.1"

/db\_xref="GI:17351923"

/translation="MGSLLPSSRRKSLPSPSSVQGGPVTMEAKSKATAVALGSPFAGPAELSLRLGLPLTVSDGWMTVLSEVSGREYNIPSVHAKVSHGLYGLSREKABELLLPGNPGGAFILRESQTRRGYSLSVRLSRPASWDRIHRHVRHCLDNGWLYISPLRTFSLQALVDHYSEGHFAPWQGYTPTCDCAEDTTQLERAGQLPPVF"

ORIGIN

Query Match 55.9%; Score 661.8; DB 9; Length 737;

Best Local Similarity 93.3%; Pred. No. 4.7e-166;

Matches 734; Conservative 0; Mismatches 2; Indels 51; Gaps 2;

Qy 398 ATGGGAATCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 457

Db 1 ATGGGAATCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 60

Qy 458 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 517

Db 61 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGGCCCTG 120

Qy 518 GCGAGTTTCCCGCAGGTGGCCCGCGAGCTGCTGCTGAGACTCGGAGGCCATTGACC 577

Db 121 GCGAGTTTCCCGCAGGTGGCCCGCGAGCTGCTGCTGAGACTCGGAGGCCATTGACC 180

Qy 578 ATCGTCTCTGAGGATGGAGACTGGTGGACCGTGTGTCTGAAGTCTCAGGAGAGATAT 637

Db 181 ATCGTCTCTGAGGATGGAGACTGGTGGACCGTGTGTCTGAAGTCTCAGGAGAGATAT 240

Qy 638 AACATCCCGAGGTCACGTCGGGCAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 697

Db 241 AACATCCCGAGGTCACGTCGGGCAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC 300

Qy 698 AGGAGAGAAAGCAGAGAGAACTGCTGTGTTTACTTGGGAAACCTTGGAGGGCTTCTCTATC 757

Db 301 AGGAGAGAAAGCAGAGAGAACTGCTGTGTTTACTTGGGAAACCTTGGAGGGCTTCTCTATC 360

Qy 758 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCTGAGTCCGCTCAGCGCCCTGCA 817

Db 361 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCTGAGTCCGCTCAGCGCCCTGCA 420

Qy 818 TCCTGGACCGGATCAGACACTCAGGATCCACTGCTTGCACCAATGGCTGTACATC 877

Db 421 TCCTGGACCGGATCAGACACTCAGGATCCACTGCTTGCACCAATGGCTGTACATC 480

Qy 878 TCACCGGCTCAGCTCCCTCCTACTCCAGGCGCTGTGAGCACTTACTCTGAGCTGGG 937

Db 481 TCACCGGCTCAGCTCCCTCCTACTCCAGGCGCTGTGAGCACTTACTCTT 531

Qy 938 GATGACATCTGCTCCTACTCAAGGAGCCCTGTGTCTCTGTCAGAGGGCTGGCCGCTCCCT 997

Db 532 -----GAGGCTGGCCGCTCCCT 550

Qy 998 GCGAAGATATACCCCTACCTGTGTCAGAGGACACCACTCACTGGAAGAGCTG 1057

Db 551 GCGAAGATATACCCCTACCTGTGTCAGAGGACACCACTCACTGGAAGAGCTG 610

Qy 1058 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAAGGT 1117

Db 611 GACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAAGGT 670

Qy 1118 CTCGGAGTCCCTCAGCTTCTATCAGCTGGAATGA-CGAGGCTGCTCTTTGGATGA 1176

Db 671 CTCGGAGTCCCTCAGCTTCTATCAGCTGGAATGA-CGAGGCTGCTCTTTGGATGA 730

Qy 1177 TGCCTAG 1183

Db 731 TGCCTAG 737

RESULT 13

LOCUS BC052655 1384 bp mRNA linear ROD 04-NOV-2003

DEFINITION Mus musculus Src-like-adaptor 2, mRNA (cDNA clone MGC:60811 IMAGE:30040401), complete cds.

## ACCESSION

BC052655

## VERSION

BC052655.1

## KEYWORDS

MGC

## SOURCE

Mus musculus

## ORGANISM

Mus musculus (house mouse)

## REFERENCE

1 (bases 1 to 1384)

## AUTHORS

Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Dietzenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, V.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

## TITLE

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

## MEDLINE

22388257

## PUBMED

12477932

## REFERENCE

2 (bases 1 to 1384)

## AUTHORS

Strausberg, R.

## TITLE

Direct Submission

## JOURNAL

Submitted (16-MAY-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

## COMMENT

Contact: MGC help desk  
Email: [cgabp-remail.nih.gov](mailto:cgabp-remail.nih.gov)  
Tissue Procurement: Drs. Dennis Taub, Dan Longo (NIA, USA), Jonathan Keller (NCI, USA)  
cDNA Library Preparation: Yulan Piao and Minoru Ko (National Institute on Aging, NIH: <http://lgsun.grc.nia.nih.gov/cDNA/>)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLUMINA)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)  
Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirip, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

## FEATURES

Location/Qualifiers

1..1384

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6Ncr"

## SOURCE

1..1384

/clone="MGC:60811 IMAGE:30040401"

/tissue\_type="Hematopoietic Stem Cell,"

(Lin-/c-Kit+/Sca-1-), mouse, 10 wks"

/clone\_lib="NIA Mouse Hematopoietic Stem Cell

(Lin-/c-Kit+/Sca-1-) cDNA Library (long)"

/lab\_host="DH10B"

/note="Vector: pSPORT1"

1..1384

/gene="Slap2"

/note="synonyms: SLAP2, SLAP-2"

/db\_xref="LocusID:77799"

/db\_xref="MGI:1925049"

330..1109

/codon\_start=1

/product="Slap2 protein"

/protein\_id="AAH52655.1"

/db\_xref="GI:30851668"

/db\_xref="LocusID:77799"

/translation="MGSLSRGKTSPPSSSGDQEPVSMOPERRKVTAVAGSPPA

GEQARLSRLGEPPLTISEDDQWTVQSEVSGREYHMSVYVAKVAGHLEGLSREK

AEELLLPENPGAFILRESQTRRCISLVRLSPASWDRHRRHRLQRLDGLWLYIS

PLRTFPLHALVEHYSELDAGICCPRLPCLQKGLPGLPGKDPPTVPTVPTSSLLAWKK

LDRLSLFLFAPASGEASLLSEGLRESLSYISLAEDPLDDA"

432..596

/note="SH3; Region: SH3 domain. SH3 (Src homology 3)

domains are often indicative of a protein involved in

signal transduction related to cytoskeletal organisation.

First described in the Src cytoplasmic tyrosine kinase.

The structure is a partly opened beta barrel"

/db\_xref="CDD:pfam00018"

603..899

/note="SH2; Region: Src homology 2 domains"

/db\_xref="CDD:cd00173"

## ORIGIN

Query Match 47.6%; Score 563; DB 10; Length 1384;

Best Local Similarity 78.3%; Pred. No. 1.3e-139;

Matches 715; Conservative 0; Mismatches 190; Indels 8; Gaps 3;

QY 274 AGGCGCCCAAGACCTTACCTGTCCAGCAGCAGCATGGTCTCAGCAGAGCTGTCTTCC 333

DB 208 AGGACCTGCAAGGCTTACCTGTGGTCACTG-TGCACATTGGCTGACTACCTCAT 265

QY 334 CAAGCCTTTGATGACAAACCAATTTCCCTCGATGATGCTTCTGAGTCTCTGTGAGG 393

DB 266 CAAACGCTGTGATGGCAACCTTTCCCTTCCAGGTTCACTGTGCTGTGAGGCTGTGCTG 325

QY 394 AACATGGGAAGTCTCCCGCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTC 453

DB 326 AGTGATGGGAAGTTTGTCTCAGCAGAGGGAACCT---CCAGCCCGCCAGCTCCTC 382

QY 454 TGTCCAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAACAGGCCACAGCCGTGGC 513

DB 383 TGTCTCAGCAGCAGGAACCCGTGTCATGCAACAGAGAGCAGCAAGGTACAGCTGTGGC 442

QY 514 CTGGGCACTTTCCCGCAGGTGGCCCGCCAGCTGTGCTGAGACTCGGGAGGCCATT 573

DB 443 CTGGGCACTTTCCCGCAGGTGGAACAGCCAGACTATCTCTGAGACTCGGGAGGCCGT 502

QY 574 GACCATCGTCTCTGAGGATGGAGACTGTGTGGACGGTGTCTGTAAGTCTCAGGCAGAGA 633

DB 503 GACCATCATCTCTGAGGATGGAGATTGTGTGGACAGTCCAGTCCGAAGTCTCAGGCAGAGA 562

QY 634 GTATACATCCCGAGCTCCAGTGGGCAAGTCTCCCATGAGTGTGCTGTATGAGGCCCT 693

DB 563 GTACCATATGCCCGAGTGTGTATGTGGCTAAAGTCCGCCACGGGTGGCTGTACGAGGCCCT 622

QY 694 GAGCAGGAGAAAGCAGAGGAACTGTGTGTGTACTGGGAACTCTGGAGGGGCTTCTCT 753

DB 623 GAGCCGGGAGAAAGCCGAGAACTACTCTGTGTACTGGGAACTCTGGAGGGGCTTCTCT 682

QY 754 CATCCGGGAGAGCCAGACACAGGAGAGGCTCTTTACTCTCTGTCTGAGTCCGCTCAGCGGCC 813

```
Db 683 CATCCGGAGAGCCAGACACAGGAGAGCTGCTATTCCCTCTCCGTCCGACTCAGCCGCC 742
Qy 814 TGCATCTCTGGACCGGATCAGACACTACAGGATCCACTGCTTTGACAAATGGCTGGCTGTA 873
Db 743 TGCATCTTGGACCGGATCAGACACTACAGGATACAGCGTCTTTGACAAATGGCTGGCTGTA 802
Qy 874 CATCTCACCGGCTCCTCCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 933
Db 803 CATCTCACCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 862
Qy 934 GCGGATCAGATCTGCTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 993
Db 863 AGCAGATGGCATCTGCTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 922
Qy 994 CCCTGGCAAGATATACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1053
Db 923 ACCTGGCAAGATACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 982
Qy 1054 GCTGCAAGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1110
Db 983 GCTGCAAGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1042
Qy 1111 TGAGGCTCTCGGAGTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1170
Db 1043 TGAGGCTCTCGGAGTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1102
Qy 1171 GGATGATGCTCTAG 1183
Db 1103 TGCTTAGCCCTGG 1115

RESULT 14
AX511150
LOCUS AX511150 1348 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1 from Patent WO0242452.
ACCESSION AX511150
VERSION AX511150.1 GI:23392044
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Meglade, J.C. and Loreto, M.P.
JOURNAL Adapter gene
JOURNAL Patent: WO 0242452-A 1 30-MAY-2002;
JOURNAL The Hospital for Sick Children (CA)
FEATURES
source
1. 1348
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 47.5%; Score 561.4; DB 6; Length 1348;
Best Local Similarity 78.2%; Pred. No. 3.6e-139;
Matches 714; Conservative 0; Mismatches 191; Indels 8; Gaps 3;

Qy 274 AGGCCCCCAAGCCCTTAACCTGCTCCAGCCAGAGCATGCTCTCAGCAGAGCTGTCTTCC 333
Db 160 AGGACCTGCAAGGCTGACCTGCTGGGTCAGTG--TGACATTTGGCTGACTACCCCTCAT 217
Qy 334 CAAGCTTTGATGACAAACCAATTTCCCTCGATGATGCTTCTGAGTGCTCTGCTGAGG 393
Db 218 CAAGCTTGATGACAAACCTTTCCCTTCCAGTTCAAGTTCAAGTTCAAGTTCAAGTTCAAGTT 277
Qy 394 AACATGGGAAGTGTGCCACAGAGAAATCTCTCCAAAGCCCAAGCTTGTAGTTCTCCTC 453
Db 278 AGTGATGGGAAGTGTGCCACAGAGAAACCT---CCAGCCCCAGCCCCAGCTCCTCCTC 334
Qy 454 TGTCAGGCCAGGACCTGTGACATGGAAGCAGAGAGCAAGCCACAGCCGCTGGC 513
```

```
Db 335 TGGTCCAGCAGGNAACCCGTGTCTCCATCCACACAGAAAGACACAAGTFCACAGCTGTGGC 394
Qy 514 CTTGGGCAAGTTTCCCGCAGAGTGGCCCGGCGAGCTGTGCTGAGACTCAGCGGAGCCATT 573
Db 395 CTTGGGCAAGTTTCCCGCAGAGTGAACAGCGAGACTATCTCTGAGACTCAGCGGAGCCGCT 454
Qy 574 GACCATCTCTCTGAGGATGAGAGCTGGTGACAGCTGTCTGCTGAAGTCTCAGGACAGAGA 633
Db 455 GACCATCATCTCTGAGGATGAGAGTTGGTGACAGTCCAGTCGGAGTCTCAGGACAGAGA 514
Qy 634 GTATAACATCCCGCAGCTCCACGTTGGGCAAAAGTCTCCCATGGGTGGCTGTATGAGGGCT 693
Db 515 GTATCACATATGCCCACTGTGTATGTGGCTAAAGTCCGCCACGCGTGGTGTATGAGGGCT 574
Qy 694 GAGCAGGAGAGAAAGCAGAGGAAGTGTCTTGTACCTGGGAACCTCTGAGGAGGCTTCTCT 753
Db 575 GAGCAGGAGAGAAAGCAGAGGAAGTGTCTTGTACCTGGGAACCTCTGAGGAGGCTTCTCT 634
Qy 754 CATCCGGAGAGCCACAGCAGAGAGGCTCTTACTCTCTGTCACTCCAGCCCTCAGCGGCC 813
Db 635 CATCCGGAGAGCCACAGCAGAGAGGCTGTATTCTCTCCGTCCGACTCAGCGGCC 694
Qy 814 TGATCTCTGGGACCGGATCAGACACTACAGATCCACTGCTTGCATGCTGCTGTA 873
Db 695 TGATCTCTGGGACCGGATCAGACACTACAGATCAGCGTCTTGCATGCTGCTGTA 754
Qy 874 CATCTCACCGGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 933
Db 755 CATCTCACCTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 814
Qy 934 GCGGATGACATCTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 993
Db 815 AGCAGATGGCATCTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 874
Qy 994 CCTCGCAAGGATATACCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1053
Db 875 ACCTGGCAAGATACACCTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 934
Qy 1054 GCTGACAGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1110
Db 935 GCTGACCGCAGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 994
Qy 1111 TGAGGCTCTCGGAGTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1170
Db 995 TGAGGCTCTCGGAGTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1054
Qy 1171 GGATGATGCTCTAG 1183
Db 1055 TGCTTAGCCCTGG 1067

RESULT 15
AF287467
LOCUS AF287467 1348 bp mRNA linear ROD 03-JUN-2002
DEFINITION Mus musculus Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF287467
VERSION AF287467.1 GI:17351918
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Loreto, M.P., Berry, D.M. and McGlade, C.J.
JOURNAL Functional cooperation between c-Cbl and Src-like adaptor protein 2
MEDLINE in the negative regulation of T-cell receptor signaling
PUBMED Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
AUTHORS 22022020
JOURNAL 2 (bases 1 to 1348)
REFERENCE 12024036
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for
```

Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada

FEATURES

source

Location/Qualifiers

1..1348

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/dev\_stage="day 15 embryo"

282..1081

/note="SLAP-2"

/codon\_start=1

/product="Src-like adaptor protein-2"

/protein\_id="AAL38196.1"

/db\_xref="GI:17351919"

/translation="MGSLSRGTSSPPSSSPDQEPVSMOPERHKVTAVALGSPFA

GFQARLSRLGPLELRIISDGDWMTVQSEVGRVHPSVXVAKYAHQWLEGLSREK

ABELLLPNPGAGLIRDSQTRGCVSLSVLSPASWDRIRHVRIRLQNLGWLVIY

PLITPSLHALVEHSELADGCCIRFCVLPKLPFGKDPFPVPTVPTSSLNWKK

LDRSLFLFAPASGEASLSEGLRESLSYSLSLAEDPLDDA"

381..551

/note="Region: SH3 domain"

552..824

/note="Region: SH2 domain"

1301..1306

ORIGIN

Query Match 47.5%; Score 561.4; DB 10; Length 1348;

Best Local Similarity 78.2%; Pred. No. 3.6e-139;

Matches 714; Conservative 0; Mismatches 191; Indels 8; Gaps 3;

QY	274	AGGGCCCCCAAGCCCTACCTGCTCCAGCAGAGATCGGTCCAGCAGAGCTGTCCTCC	333
Db	160	AGGACCTGCAAGGCCCTGACCTGTCGGGTCACTG--TGACATTTGGCTGACTACCTCAT	217
QY	334	CAAGCCCTTTGATGACAAACCAATTTCCCTCGATGATGTCCTTCTGAGTGCTCTGCTGAGG	393
Db	218	CAAACTCTGATGGCAAACTTTCCCTTTCCAGGTTCACTGTCGTTGTGAGCGTCTGCTG	277
QY	394	AACATGGGAAGTCTGCCAGCAGAGAAATCTTGCCAAAGCCCAAGCTTGAGTTCTTC	453
Db	278	AGTGATGGGAAGTTTTCAGCAGAGGGAACCT---CCAGCCCCAGCCCCAGCTCTCTC	334
QY	454	TGTCGAAGCCCAAGGACCTGTGACCATGGAAGCAGAGAGAAACAAAGGCCACAGCCGTGGC	513
Db	335	TGTTCCAGACACAGAACCCGTTGTCATGCAACAGAAAGACACAGGTTCACAGCTGTGGC	394
QY	514	CCTGGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCGTGAGACTCGGGAGGCCATT	573
Db	395	CCTGGCAGTTTCCCGCAGGTGAACAGGCCAGACTATCTCTGAGACTCGGGAGCCGCT	454
QY	574	GACCATCTCTCTGAGGATGGAGACTGGTGACGGTGTCTGCTGAAGTCTCAGGCAGAGA	633
Db	455	GACCATCTCTCTGAGGATGGAGATTGGTGGACAGTCCAGTCGGAAGTCTCAGGCAGAGA	514
QY	634	GTATAACATCCCGCAGCGTCCAGTGGGCAAGTCTCCATGGGTGGCTGTATGAGGGCCT	693
Db	515	GTACCACATGCCCGCAGTGTATGTGGCTAAAGTCGCCACCGGTGGCTGTACGAGGGCCT	574
QY	694	GAGCAGGAGAGAGCAGAGAACTGCTGTGTGTACTGGACCCCTGGAGGGCCCTTCCT	753
Db	575	GAGCAGGAGAGAGCAGAGAACTGCTGTGTGTACTGGACCCCTGGAGGGCCCTTCCT	634
QY	754	CATCCGGGAGAGCCAGACAGGAGGCTTTTACTCTCTGTGTCAGTCCGCTCAGCCGCC	813
Db	635	CATCCGGGAGAGCCAGACAGGAGGCTGTATTCCCTGTCCGCTCCGACTCAGCCGCC	694
QY	814	TGCATCTGGGACCGGATCAGACACTACAGGATCCACTGCCTTGACATGGCTGGCTGTA	873
Db	695	TGCATCTGGGACCGGATCAGACACTACAGGATCAGCGCTCTTGACAATGGCTGGCTGTA	754
QY	874	CATCTCAGCCGCGCTCACCTTCCCTCTCACTCCAGGCCCTGGTGGACCATTTACTCTGAGCT	933
Db	755	CATCTCAGCTCGCTCACCTTCCCTCTCACTCCAGGCCCTGGTGGAGCATTTACTCTGAGCT	814

Search completed: February 19, 2004, 23:23:13  
Job time : 5187.29 secs

QY	934	GGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTGAGAGGCTGGCCCGCT	993
Db	815	AGCAGATGGCATCTGCTGTCTCCCTCAGGGAGCCGCTGTCTCTGAGAAAGCTTGGCCACT	874
QY	994	CCCTGGCAAGGATATACCCCTTACTGTGACTGTGACAGGAGACACCACTCAACTGGAAGA	1053
Db	875	ACCTGGCAAGGATATACACCTTCACTGTGACTGTGCCACATCATCACTAAATTTGGAATA	934
QY	1054	GCTGGACAGCTCCCTCTCTGTTTTCTGAAG---CTGCCACAGGGAGGAGTCTCTTCTCAG	1110
Db	935	GCTGGACCGAGCCTCTGTTTTCTGGAAGCACCTGCCAGTGGGGAGGCATCTCTGCTCAG	994
QY	1111	TGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCTTCTGCTTCTTCTCTTTT	1170
Db	995	TGAGGGCTCCGAGAGTCCCTCAGTTCCTTACATGAGCTGGCTGGTGGAGACCCCTTGGATGA	1054
QY	1171	GGATGATGCTAG 1183	
Db	1055	TGCTTAGCCCTGG 1067	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:15:48 ; Search time 1096 Seconds  
(without alignments)  
4585.415 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccaaggacc.....tctcttggatgcctag 1183

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N: Geneseq\_25Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1183	100.0	1183	6	ABK61465 Human cDN
2	1181.4	99.9	2567	6	AD43980 Human Src
3	784.4	66.3	786	6	AA144089 Human mod
4	784.4	66.3	786	6	ABQ74343 Human Src
5	758.2	64.1	763	6	ABQ98670 Human ORF
6	740.2	62.6	837	3	AACT7202 Human ORF
7	724.2	61.2	1413	6	ABQ99374 Human cod
8	661.8	55.9	737	6	AA144090 Mouse MAR
9	651.2	55.0	2049	5	AA574750 DNA encod
10	561.4	47.5	1348	6	AA144087 Mouse mod
11	405	34.2	603	5	AA574748 DNA encod
12	395	33.4	445	5	AA574747 DNA encod
13	348	29.4	444	6	ABQ98669 Human ORF
14	341	28.8	875	6	ABQ99151 Human ORF
15	191.2	16.2	211	5	AA570181 DNA encod
16	158.6	13.4	2109	4	AA502049 DNA encod
17	157.4	13.3	2665	6	ABL65189 Lung canc
18	157.4	13.3	2665	6	ABK83738 Human cDN
19	157.4	13.3	3756	7	ACC81091 Human Src
20	157.4	13.3	3756	7	ABX62975 Human act
21	141.8	12.0	432	4	AA112879 Probe #28
22	141.8	12.0	432	4	ABA54580 Human foe
23	141.8	12.0	432	4	AA134236 Probe #29

24	141.8	12.0	432	4	ABA44128 Human bre
25	141.8	12.0	432	4	ABA24363 Probe #28
26	141.8	12.0	432	4	AAK28314 Human bon
27	141.8	12.0	432	4	AAK02872 Human bra
28	141.8	12.0	432	4	ABS27912 Human liv
29	141.8	12.0	432	5	AAI02797 Probe #27
30	141.8	12.0	432	5	ABS02823 Human gen
31	141.8	12.0	432	4	AAI14520 Probe #44
32	141.8	12.0	448	4	AAI14520 Human gen
33	123	10.4	1926	6	ABS04499 Human gen
34	123	10.4	2015	6	ABK83940 Human cDN
35	123	10.4	2015	6	ABL66673 Lung canc
36	123	10.4	2015	6	ABK83939 Human cDN
37	123	10.4	2015	9	AA621155 Human hae
38	119.4	10.1	2341	9	AD553793 Human pro
39	119.4	10.1	2288	6	ABK83935 Human cDN
40	110.6	9.3	1530	8	ADA02968 Human LCK
41	110.6	9.3	1530	9	ADB72706 Human LCK
42	110.6	9.3	1530	9	ADC85448 Human LCK
43	110.6	9.3	2032	3	AA246491 PKA subst
44	110.6	9.3	2032	7	ACC72850 Human can
45	110.6	9.3	2032	8	ADA02967 Human LCK

ALIGNMENTS

RESULT 1

ABK61465

ID ABK61465 standard; cDNA; 1183 BP.

XX AC ABK61465;

XX DT 18-JUN-2002 (first entry)

XX DE Human cDNA encoding protein NOV13.

XX DX Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;

KW cell signal processing disorder; metabolic pathway modulation disorder;

KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;

KW uterine cancer; immune response; graft-versus-host disease;

KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;

KW hypertension; congenital heart defects; multiple sclerosis; inflammation;

KW Albright hereditary osteodystrophy.

XX OS Homo sapiens.

XX PN WC200216599-A2.

XX PD 28-FEB-2002.

XX PF 27-AUG-2001; 2001WO-US026510.

XX PR 25-AUG-2000; 2000US-0228191P.

XX PR 08-FEB-2001; 2001US-0267300P.

XX PR 20-FEB-2001; 2001US-0269961P.

XX PR 20-MAR-2001; 2001US-0277337P.

XX PA (CURA-) CURAGEN CORP

XX PA (COR-) COR THERAPEUTICS INC.

XX PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;

XX PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;

XX DR WPI; 2002-280937/32.

XX DR P-PSDB; AAU91308.

XX PT New polypeptides for treating or preventing a disorder associated with

XX PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.

XX PS Claim 1; Page 98; 263pp; English.

XX CC The invention relates to an isolated polypeptide (NOVX) a mature form of

CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide  
 CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
 CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it  
 CC and antibody against it, are useful for treating or preventing (e.g. by  
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOVX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation or Albrit  
 CC hereditary osteodystrophy, and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC for developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence encodes a NOVX protein  
 XX  
 SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Query Match 100.0%; Score 1183; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-310;  
 Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTAGAGTCCAAAGACCCACGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
 DB 1 AGCTAGAGTCCAAAGACCCACGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
 QY 61 CCTTCCCTCCCTGGCTGTGCTGGAGGGTTCCTCCAGTCCAGAACTCCCTAAGGAG 120  
 DB 61 CCTTCCCTCCCTGGCTGTGCTGGAGGGTTCCTCCAGTCCAGAACTCCCTAAGGAG 120  
 QY 121 CATGGGAGCTGATCCATCCTGTGTGTACAACTGTGCTGACGAGAGATGCTGAGCT 180  
 DB 121 CATGGGAGCTGATCCATCCTGTGTGTACAACTGTGCTGACGAGAGATGCTGAGCT 180  
 QY 181 ACCCAAAACCAACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAGTCTGGGTG 240  
 DB 181 ACCCAAAACCAACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAGTCTGGGTG 240  
 QY 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCAAGCCCTAACCTGTCCA 300  
 DB 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCAAGCCCTAACCTGTCCA 300  
 QY 301 GCCAGAGCATGCTCTCAGCAGAGCTGTCTCCCAAGCTTTGATGACAAACCAATTTCC 360  
 DB 301 GCCAGAGCATGCTCTCAGCAGAGCTGTCTCCCAAGCTTTGATGACAAACCAATTTCC 360  
 QY 361 CTCGATGATGCTCTTCTGAGTCTCTGCTGAGGAACTGCTGCCAGCAGAG 420  
 DB 361 CTCGATGATGCTCTTCTGAGTCTCTGCTGAGGAACTGCTGCCAGCAGAG 420  
 QY 421 AAAATCTCTGCCAAGCCCAAGCTTCAAGTCTCTGCTGACAGGAGCTGTGACCAT 480  
 DB 421 AAAATCTCTGCCAAGCCCAAGCTTCAAGTCTCTGCTGACAGGAGCTGTGACCAT 480  
 QY 481 GGAAGCAGAGAGAAAGGCCACAGCCGTGGCCCTGGGCGAGTTTCCCGCAGGTGGCCC 540  
 DB 481 GGAAGCAGAGAGAAAGGCCACAGCCGTGGCCCTGGGCGAGTTTCCCGCAGGTGGCCC 540  
 QY 541 GGCAGAGCTGTGCTGAGACTGCGGGAGCCATTGACCATCTCTCTGAGGATGAGAGCTG 600  
 DB 541 GGCAGAGCTGTGCTGAGACTGCGGGAGCCATTGACCATCTCTCTGAGGATGAGAGCTG 600  
 QY 601 GTGGACGGTGTCTGTGAAGTCTCAGGAGCAGAGATTAACATCCCGAGCTCCACGTGGG 660  
 DB 601 GTGGACGGTGTCTGTGAAGTCTCAGGAGCAGAGATTAACATCCCGAGCTCCACGTGGG 660

QY 661 CAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGTAGCGGGGAGAAAGCAGAGAACTGCT 720  
 DB 661 CAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGTAGCGGGGAGAAAGCAGAGAACTGCT 720  
 QY 721 GTTGTACTCTGGAAACCCCTGGAGGGCCCTTCTCATCCGGGAGACGACAGACGAGAGG 780  
 DB 721 GTTGTACTCTGGAAACCCCTGGAGGGCCCTTCTCATCCGGGAGACGACAGACGAGAGG 780  
 QY 781 CTCTTACTCTCTGTAGTCCGGCTCAGCGCCCTGATCTCTGGGACCGGATCAGACACTA 840  
 DB 781 CTCTTACTCTCTGTAGTCCGGCTCAGCGCCCTGATCTCTGGGACCGGATCAGACACTA 840  
 QY 841 CAGGATCCACTGCTCTGTGACAAATGGCTGTGTACATCTCACCGCGCTCACCTTCCCTTC 900  
 DB 841 CAGGATCCACTGCTCTGTGACAAATGGCTGTGTACATCTCACCGCGCTCACCTTCCCTTC 900  
 QY 901 ACTCCAGGCTGTGGGACCACTTACTCTGAGTGGCGGATGACATCTGCTGCTACTCAA 960  
 DB 901 ACTCCAGGCTGTGGGACCACTTACTCTGAGTGGCGGATGACATCTGCTGCTACTCAA 960  
 QY 961 GGAGCCCTGTCTCTGACAGAGGGCTGGCCCGCTCCCTGCAAGGATATACCCCTACTCT 1020  
 DB 961 GGAGCCCTGTCTCTGACAGAGGGCTGGCCCGCTCCCTGCAAGGATATACCCCTACTCT 1020  
 QY 1021 GACTGTGACAGAGACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGA 1080  
 DB 1021 GACTGTGACAGAGACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGA 1080  
 QY 1081 AGCTGCCACAGGGGAGGAGTCTTCTCAGTCAAGGCTCTCCGGGAGTCCCTCAGCTTCTA 1140  
 DB 1081 AGCTGCCACAGGGGAGGAGTCTTCTCAGTCAAGGCTCTCCGGGAGTCCCTCAGCTTCTA 1140  
 QY 1141 CATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTTAG 1183  
 DB 1141 CATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTTAG 1183

RESULT 2

AAD43980  
 ID AAD43980 standard; cDNA; 2567 BP.

XX AAD43980;

XX 13-DEC-2002 (first entry)

XX Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.

XX Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 CDS 415..1200

XX FT /\*tag= a

XX FT /product= "Human SLAP-2"

XX WO200242457-A1.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US043367.

XX 22-NOV-2000; 2000US-0252545P.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;

WPI; 2002-463632/49.  
P-PSDB; AAE26357.

Novel substantially purified human SH2/SH3-domain-containing adapter polypeptide, termed Src-like Adapter Protein-2, useful for therapeutic intervention in immunological and inflammatory disorders and cancer.

Claim 2; Fig 1; 85pp; English.

The invention relates to a substantially purified human SH2/SH3-domain-containing adapter polypeptide, termed Src-like Adapter Protein-2 (SLAP-2). The invention is useful for treating an immune disorder involving hyperactivity of B- or T- lymphocytes in a mammal. The invention is useful for screening for antagonists or inhibitors of the interaction of hSLAP-2 with cellular signalling compounds, for diagnosing, treating or preventing diseases or disorders associated with aberrant or uncontrolled cellular signal transduction, for determining those cellular signalling molecules which associate with hSLAP-2 and which provide critical signals for cell activation, and as effectors in methods to affect T- cell activation. The invention is useful in screening assays to identify and detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for potential use to treat autoimmune diseases which may be caused by hyperactivated B cells, as well as to treat diseases which may be caused by hyperactivated T cells, in addition to other immune system related conditions, diseases, or disorders. T-cell and B-cell neoplasms, inflammation disorders, diseases and conditions, rheumatoid arthritis, osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's and ulcerative colitis), allergies, particularly those involving hyperactivity of B-cells and T- cells, or other immune cells, such as mast cells or eosinophils, autoimmune diseases such as systemic lupus erythematosus and multiple sclerosis, pulmonary diseases including asthma, acute respiratory distress syndrome, and chronic obstructive pulmonary disorder, tissue/ organ rejection and cancer. The invention is useful in gene therapy. The present sequence is human SLAP-2 cDNA

Sequence 2567 BP: 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

481	Qy	481	GGAAAGCAGAGAGAAAGCAAGGCCACACCGCTGGCCCTGGGCAAGTTCCTCCGGCAGGTGGCCC	540
498	Db	498	GGAAAGCAGAGAGAAAGCAAGGCCACACCGCTGGCCCTGGGCAAGTTCCTCCGGCAGGTGGCCC	557
541	Qy	541	GGCCGAGCTGTCTCGCTGAGACTCGGGGAGGCCATTGACCATCTGTCTCTGAGGATGGAGACTG	600
558	Db	558	GGCCGAGCTGTCTCGCTGAGACTCGGGGAGGCCATTGACCATCTGTCTCTGAGGATGGAGACTG	617
601	Qy	601	GTGGACGGTGTCTGTCTGAAGTCTCAGGCGAGAGAGTATTAACATCCCGAGCTCCACGTGGG	660
618	Db	618	GTGGACGGTGTCTGTCTGAAGTCTCAGGCGAGAGAGTATTAACATCCCGAGCTCCACGTGGG	677
661	Qy	661	CAAAAGTCTCCATGAGTGGTGTATCAGGCGCTTGAGCAGGAGCAAAAGCAGAGGAACTGTCT	720
678	Db	678	CAAAAGTCTCCATGAGTGGTGTATCAGGCGCTTGAGCAGGAGCAAAAGCAGAGGAACTGTCT	737
721	Qy	721	GTGTGTACTGTGGAAACCTGTGGAGGGGCTTCTCTCATCCCGGAGAGCGAGACACGAGGAGG	780
738	Db	738	GTGTGTACTGTGGAAACCTGTGGAGGGGCTTCTCTCATCCCGGAGAGCGAGACACGAGGAGG	797
781	Qy	781	CTCTTACTCTCTGTAGTCCGCTCAGCGCCCTGCATCCTGTGGAGCCGGATCAGACACTA	840
798	Db	798	CTCTTACTCTCTGTAGTCCGCTCAGCGCCCTGCATCCTGTGGAGCCGGATCAGACACTA	857
841	Qy	841	CAGGATCCACTGTCCTTGACAAATGGCTGGCTGTACATCTCACCGCGGCTCACCTTCCGCTC	900
858	Db	858	CAGGATCCACTGTCCTTGACAAATGGCTGGCTGTACATCTCACCGCGGCTCACCTTCCGCTC	917
901	Qy	901	ACTCCAGGCCCTGGTGGACCAATTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCTAA	960
918	Db	918	ACTCCAGGCCCTGGTGGACCAATTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCTAA	977
961	Qy	961	GGAGCCCTGTGTCTCTGCAGAGGGCTGGCCCGCTCCCTCTGGCAAGGATATACCCCTACCTGT	1020
978	Db	978	GGAGCCCTGTGTCTCTGCAGAGGGCTGGCCCGCTCCCTCTGGCAAGGATATACCCCTACCTGT	1037
1021	Qy	1021	GACTGTGCAGAGACACCACTCAACTGGAAAGAGCTGGAGCAGCTCCCTCTCTGTTTCTGA	1080
1038	Db	1038	GACTGTGCAGAGACACCACTCAACTGGAAAGAGCTGGAGCAGCTCCCTCTCTGTTTCTGA	1097
1081	Qy	1081	AGTGTGCCACAGGGGAGGAGTCTCTTCTCAGTGAAGGCTCTCCGGGAGTCCCTCAGCTTCTA	1140
1098	Db	1098	AGTGTGCCACAGGGGAGGAGTCTCTTCTCAGTGAAGGCTCTCCGGGAGTCCCTCAGCTTCTA	1157
1141	Qy	1141	CATCAGCTGAATGACAGGCTGTCTCTTTGGATGATGCCTAG	1183
1158	Db	1158	CATCAGCTGAATGACAGGCTGTCTCTTTGGATGATGCCTAG	1200

RESULT 3  
AAL44089

AAL44089  
 ID AAL44089 standard: cDNA: 786 bp.

XX  
AC AAL44089;

XX DT 03-OCT-2002 (first entry)

XX DE Human mobility at the onset of the COVID-19 pandemic

100

Human; gene; ss; gene therapy; modulator of antigen receptor signalling;  
MARS; tumour suppressor gene; Src-like adapter protein; SLAP;  
myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
immunosuppression; myeloproliferative disorder; breast cancer.

OS Homo sapiens.

XX	Key	Location/Qualifiers
EH		

FT	CDS	1. .786

```
ET / *tag= a
ET / *tag= a
```

... THE PRODUCTION OF HUMAN PROTEIN

PN WO200242452-A2

XX 30-MAY-2002.  
 XX 26-NOV-2001; 2001WO-CA001662.  
 XX 27-NOV-2000; 2000CA-02324663.  
 XX (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 XX Mcglade JC, Loreto MP;  
 XX WPI; 2002-566564/60.  
 XX P-PSDB; AAO15457.  
 XX New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX Claim 12; Page 75; 110pp; English.  
 XX The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present cDNA sequence encodes a human MARS protein  
 XX  
 XX Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;  
 Query Match 66.3%; Score 784.4; DB 6; Length 786;  
 Best Local Similarity 99.9%; Pred. No. 1.9e-202;  
 Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTCCCAAGCCCAAGCTTGATTCCTCTGTC 457  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTCCCAAGCCCAAGCTTGATTCCTCTGTC 60  
 QY 458 CAAGGCCAGGGACCTGTGACCATGGGAAGCAGAGAGAGCAAGCCACGCTGGCCCTG 517  
 Db 61 CAAGGCCAGGGACCTGTGACCATGGGAAGCAGAGAGAGCAAGCCACGCTGGCCCTG 120  
 QY 518 GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGCTCGCTGAGACTCGGGAGCCATTGACC 577  
 Db 121 GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGCTCGCTGAGACTCGGGAGCCATTGACC 180  
 QY 578 ATCGTCTCTGAGATGAGACTGCTGAGCGTGTCTGTGAAGTCTCAGGCAGAGATAT 637  
 Db 181 ATCGTCTCTGAGATGAGACTGCTGAGCGTGTCTGTGAAGTCTCAGGCAGAGATAT 240  
 QY 638 AACATCCCCAGCTCCACGTTGGCCAAAGTCTCCCATGGGTGCTGTATCAGGGCCCTGAGC 697  
 Db 241 AACATCCCCAGCTCCACGTTGGCCAAAGTCTCCCATGGGTGCTGTATCAGGGCCCTGAGC 300  
 QY 698 AGGAGAAACAGAGAACTGCTGTTTACCTGGGAACCTCGGAGGCCCTTCCTCATC 757  
 Db 301 AGGAGAAACAGAGAACTGCTGTTTACCTGGGAACCTCGGAGGCCCTTCCTCATC 360  
 QY 758 CGGAGAGCCAGACACGAGAGGCTCTTACTCTCTGCTCAGTCCGCTCAGCCGCCCTGCA 817  
 Db 361 CGGAGAGCCAGACACGAGAGGCTCTTACTCTCTGCTCAGTCCGCTCAGCCGCCCTGCA 420  
 QY 818 TCCTGGAGCCGATCAGACACTACAGATCCATGCTGACATGGCTGGCTGTACATC 877  
 Db 421 TCCTGGAGCCGATCAGACACTACAGATCCATGCTGACATGGCTGGCTGTACATC 480  
 QY 878 TCACCGGCTCTACCTTCCCTCACTCCAGGCCCTGTGTGACCATTAATCTAGCTGGGG 937  
 Db 481 TCACCGGCTCTACCTTCCCTCACTCCAGGCCCTGTGTGACCATTAATCTAGCTGGGG 540  
 QY 938 GATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTGACAGAGGCTGGCCGCTCCCT 997

Db 541 GATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTCTGAGAGGCTGGCCGCTCCCT 600  
 QY 998 GGCAGGATATACCCCTACCTGTGACTGTGACAGAGGACACCACTCACTGGAAGAGCTG 1057  
 Db 601 GGCAGGATATACCCCTACCTGTGACTGTGACAGAGGACACCACTCACTGGAAGAGCTG 660  
 QY 1058 GACAGTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTTCTCAGTGAGGT 1117  
 Db 661 GACAGTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTTCTCAGTGAGGT 720  
 QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGTAATGACGAGGCTGTCTTTGGATGAT 1177  
 Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGTAATGACGAGGCTGTCTTTGGATGAT 780  
 QY 1178 GCCTAG 1183  
 Db 781 GCCTAG 786  
 RESULT 4  
 ABQ74343  
 ID ABQ74343 standard; cDNA; 786 BP.  
 XX  
 AC ABQ74343;  
 XX  
 DT 15-OCT-2002 (first entry)  
 XX  
 DE Human Src-like inhibitory molecule (SLIM) encoding cDNA.  
 XX  
 KW Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;  
 KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;  
 KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;  
 KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
 KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..786  
 FT /\*tag= a  
 FT /product= "SLIM"  
 FT /note= "Src-like inhibitory molecule"  
 XX  
 PN WO200255707-A2.  
 XX 18-JUL-2002.  
 PD  
 XX 10-JAN-2002; 2002WO-US000718.  
 PF  
 XX 10-JAN-2001; 2001US-0260953P.  
 PR  
 XX (RIGE-) RIGEL PHARM INC.  
 PA  
 XX Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
 XX WPI; 2002-575432/61.  
 DR P-PSDB; ABP52187.  
 XX  
 PT New src-like inhibitory molecule protein, useful for treating  
 PT immunodeficiency disorders and inflammatory disorders, comprises N-  
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.  
 XX  
 PS Claim 3; Fig 2A; 91pp; English.  
 CC  
 CC The present sequence encodes the human Src-like inhibitory molecule  
 CC (SLIM) protein (I). The present invention describes a SLIM protein  
 CC comprising an N-terminal myristylation sequence, an N-terminal SH2  
 CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising  
 CC an N-terminal myristylation sequence and an N-terminal SH2 domain which  
 CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and

anti-HIV activities, and can be used as a modulator of lymphocyte activation, and of ubiquitination of a Cbl target protein, and in gene therapy. (i) is useful for screening a bioactive agent capable of binding to SLIM. (i) is also useful for screening a bioactive agent capable of modulating SLIM binding. (i) or its fragments is useful in the study or in the treatment of conditions which involves this function or dysregulation of SLIM protein activity, i.e. to diagnose, treat or prevent SLIM associated disorders. (i) or the polynucleotide encoding it (ii) is useful for modulating leukocyte and/or platelet activation, for modulating antigen receptor-induced signalling and activation in leukocyte and/or platelets and for modulating antigen receptor-induced signalling and activation in lymphocytes and/or mast cells. (i) or (ii) is also useful for modulating the basal activity of lymphocytes. (i) or (ii) is useful in the treatment of immunodeficiency disorders, such as acquired immunodeficiency syndrome (AIDS), for the prevention and treatment of acute inflammatory disorders, chronic inflammatory disorders, autoimmune disorder and transplant rejection

XX  
SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Query Match 66.3%; Score 784.4; DB 6; Length 786;  
Best Local Similarity 99.9%; Pred. No. 1.9e-202;  
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGAAAGTCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGTCTCTGTC 457  
DB 1 ATGGAAAGTCTGCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGTCTCTGTC 60

QY 458 CAAGCCAGGACCTGTGACCATCGAGCAGAGAGAGAGCCAGCGCGTGGCCCTG 517  
DB 61 CAAGCCAGGACCTGTGACCATCGAGCAGAGAGAGAGCCAGCGCGTGGCCCTG 120

QY 518 GCGAGTTTCCCGCAGGTGGCCGCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 577  
DB 121 GCGAGTTTCCCGCAGGTGGCCGCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180

QY 578 ATGCTCTCTGAGATGAGAGCTGTGAGCGTGTGCTGAGAGTCTCAGGAGAGATAT 637  
DB 181 ATGCTCTCTGAGATGAGAGCTGTGAGCGTGTGCTGAGAGTCTCAGGAGAGATAT 240

QY 638 AACATCCCGACGTCACGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 697  
DB 241 AACATCCCGACGTCACGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 300

QY 698 AGGAGAAAGCAGAGAACTCTGTGTTACTCGGAAACCTGGAGGGCTTCCTCATC 757  
DB 301 AGGAGAAAGCAGAGAACTCTGTGTTACTCGGAAACCTGGAGGGCTTCCTCATC 360

QY 758 CGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCTGCA 817  
DB 361 CGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCTGCA 420

QY 818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGACAAATGGCTGGCTGATAC 877  
DB 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGACAAATGGCTGGCTGATAC 480

QY 878 TCACCGGCTCACTTCCCTCACTCAGGCCCTGTGACCAATTAATCTGAGCTGGCG 937  
DB 481 TCACCGGCTCACTTCCCTCACTCAGGCCCTGTGACCAATTAATCTGAGCTGGCG 540

QY 938 GATGACATCTGCTGCTACTCAAGAGGCTGTGCTGCTGAGAGGCTGGCCGCTCCCT 997  
DB 541 GATGACATCTGCTGCTACTCAAGAGGCTGTGCTGCTGAGAGGCTGGCCGCTCCCT 600

QY 998 GGCAGAGATATACCCCTACCTGATGCTGTGAGAGGACACCCTCACTGAGAGAGCTG 1057  
DB 601 GGCAGAGATATACCCCTACCTGATGCTGTGAGAGGACACCCTCACTGAGAGAGCTG 660

QY 1058 GACAGCTCCCTCCCTGTTTCTGAGAGTCCACAGGGAGAGTCTCTTCTCAGTGAGGT 1117  
DB 661 GACAGCTCCCTCCCTGTTTCTGAGAGTCCACAGGGAGAGTCTCTTCTCAGTGAGGT 720

QY 1118 CTCGGGAGTCCCTCAGCTTCTATCATCAGCCTGAAATGACGAGGCTGTCTCTTTGGATGAT 1177

DB 721 CTCCTGGAGTCCCTCAGCTTCTTACATCAGCTGAATGAGGCTGTCTCTTTGGATGAT 780  
QY 1178 GCTAG 1183  
DB 781 GCTAG 786

RESULT 5  
ABQ98670  
ID ABQ98670 standard; DNA; 763 BP.  
XX  
AC ABQ98670;  
XX  
DT 04-NOV-2002 (first entry)  
XX  
DE Human ORF477 coding sequence.  
XX  
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;  
KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
KW cancer; cardiovascular disease; allergy; autoimmune disease;  
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
XX  
OS Homo sapiens.  
XX  
PN US2002082206-A1.  
XX  
PD 27-JUN-2002.  
XX  
PF 30-MAY-2001; 2001US-00867550.  
XX  
PR 30-MAY-2000; 2000US-0208427P.  
XX  
PA (LEAC/) LEACH M D.  
PA (MEHR/) MEHRAN F.  
PA (CONL/) CONLEY P B.  
PA (TOPP/) TOPPER J N.  
PA (LAWD/) LAW D.  
XX  
PI Leach MD, Mehran F, Conley PB, Topper JN, Law D;  
XX  
DR WPI; 2002-626554/57.  
DR P-PSDB; ABP64107.  
XX  
PT New polypeptide designated ORFX are present in human atherogenic cells and are useful to prevent and treat ORFX-associated disorders including cancer, allergy, wound healing or autoimmune, cardiovascular or inflammatory disease.  
XX  
PS Claim 2; SEQ ID NO 953; 78pp; English.  
XX  
CC The present invention relates to novel human ORFX polypeptides and their coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences were discovered in human atherogenic cells, in particular in platelets and human umbilical vein endothelial cells (HUVEC) and are expressed in many other tissues as well. Atherogenic cells are cells which have the potential to develop atherosclerotic plaques. The ORFX polypeptides and nucleic acids are useful for treating or preventing a pathological condition associated with an ORFX-associated disorder, e.g. cancer, cardiovascular disease, allergy, autoimmune disease, wound healing, blood coagulation disorders or inflammatory disorders. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at [seqdata.uspto.gov/sequence.html?DocID=20020082206](http://seqdata.uspto.gov/sequence.html?DocID=20020082206)

XX  
SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 64.1%; Score 758.2; DB 6; Length 763;  
Best Local Similarity 99.6%; Pred. No. 2.4e-195;  
Matches 760; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 113 CTAAGAGCATGGGACGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGAT 172

Db 1 CTATGAGCATGGGAGCTGATCCATCCCTGCTGTACAACTGCTGACAGCAGAT 60  
 QY 173 GCTGAGCTACCCAAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAT 232  
 Db 61 GCTGAGCTACCCAAACCAACACCTAGCTCTCCCTGAAGATCTCCAGGCTGAGAGAT 120  
 QY 233 TCTGGGTGCTCTAGGACCAAGGACATGGGAGATTTCCAGAGGGCCCCAAAGCCCTAA 292  
 Db 121 TCTGGATGCTCTAGGACCAAGGACATGGGAGATTTCCAGAGGGCCCCAAAGCCCTAA 180  
 QY 293 CTTGTCAGCCAGCAGATGCTCTAGGAGAGCTGCTTCCCAAGGCTTTGATGACAAAC 352  
 Db 181 CTTGTCAGCCAGCAGATGCTCTAGGAGAGCTGCTTCCCAAGGCTTTGATGACAAAC 240  
 QY 353 CAATTTCCCTCGATGATGCTTCTGAGTCTCTGCTGAGGACCAATGGGAAGTCTGCC 412  
 Db 241 CAATTTCCCTCGATGATGCTTCTGAGTCTCTGCTGAGGACCAATGGGAAGTCTGCC 300  
 QY 413 AGCAGAGAAATCTCTGCCAGCCCAAGCTTCTGAGTCTCTGCTGCAAGGCCAGGACCT 472  
 Db 301 AGCAGAGAAATCTCTGCCAGCCCAAGCTTCTGAGTCTCTGCTGCAAGGCCAGGACCT 360  
 QY 473 GTGACCATGGAAGCAGAGAGAAAGCCACAGCCCTGCTGCGGAGTCTTCCCGGCA 532  
 Db 361 GTGACCATGGAAGCAGAGAGAAAGCCACAGCCCTGCTGCGGAGTCTTCCCGGCA 420  
 QY 533 GGTGGCCCGCCGAGCTGCTGCTGAGACTCGGGAGCCATGACCATGCTCTGAGGAT 592  
 Db 421 GGTGGCCCGCCGAGCTGCTGCTGAGACTCGGGAGCCATGACCATGCTCTGAGGAT 480  
 QY 593 GGAGACTGTGGAGCGTGTCTGAGTCTCAGGAGAGAGTATTAACATCCCCAGCGTC 652  
 Db 481 GGAGACTGTGGAGCGTGTCTGAGTCTCAGGAGAGAGTATTAACATCCCCAGCGTC 540  
 QY 653 CAGTGGCCAAAGTCTCCATGGTGGCTGTATGAGGGCTGAGCAGGAGAGAAAGCAG 712  
 Db 541 CAGTGGCCAAAGTCTCCATGGTGGCTGTATGAGGGCTGAGCAGGAGAGAAAGCAG 600  
 QY 713 GAACCTGCTGTGTTACTCTGGAAACCTCTGAGGGGCTTCTCATCCGGGAGAGCCAGAC 772  
 Db 601 GAACCTGCTGTGTTACTCTGGAAACCTCTGAGGGGCTTCTCATCCGGGAGAGCCAGAC 660  
 QY 773 AGGAGAGCTTACTCTGCTGATCCGCTCAGCGGCTGATCCCTGAGGAGCCAGAC 832  
 Db 661 AGGAGAGCTTACTCTGCTGATCCGCTCAGCGGCTGATCCCTGAGGAGCCAGAC 720  
 QY 833 AGACACTACAGGATCCACTCCCTTGACAATGGCTGCTGTACA 875  
 Db 721 AGACACTACAGGATCCACTCCCTTGACAATGGCTGCTGTACA 763

RESULT 6  
 AAC77202  
 ID AAC77202 standard; cDNA; 837 BP.  
 XX  
 AC AAC77202;  
 DT 08-FEB-2001 (first entry)  
 DE Human ORF2757 polynucleotide sequence SEQ ID NO:5513.  
 KW Human; open reading frame; ORF; detection; cytostatic; hepatotropic;  
 KW vulnary; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disease; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.  
 OS Homo sapiens.  
 XX WO200058473-A2.  
 FN 05-OCT-2000.  
 XX 31-MAR-2000; 2000WO-US008621.  
 XX 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX (CURA-) CURAGEN CORP.  
 XX Shimkets RA, Leach M;  
 XX WPI; 2000-602362/57.  
 DR P-PSDB; AAB42993.  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 PS Claim 5; Page 4692-4693; 5507pp; English.

CC AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORF open reading frames 1 to 3161. The ORF  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 CC antipsoriatic; antiparkinsonian; neurotropic; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antinflammatory; antibacterial; The  
 CC antiviral; antirheumatic; antithyroid; and antianaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORF-associated disorder. The nucleic acids can be used to express ORF  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX Sequence 837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;  
 Query Match 62.6%; Score 740.2; DB 3; Length 837;  
 Best Local Similarity 99.6%; Pred. No. 1.9e-190;  
 Matches 742; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 439 AAGCTTGAGTTCCTCTGCTCCAGGCCAGGACCTGTCACCATGGAAGCAGAGAGCAAA 498  
 Db 3 AAGCTTGAGTTCCTCTGCTCCAGGCCAGGACCTGTCACCATGGAAGCAGAGAGCAAA 62  
 QY 499 GGCACACGCCGCGCCCTGCGGACGTTTCCCGGACGAGTGGCCCGCGAGTGTGCTGAG 558  
 Db 63 GGCACACGCCGCGCCCTGCGGACGTTTCCCGGACGAGTGGCCCGCGAGTGTGCTGAG 122  
 QY 559 ACTCGGGGAGCATTGACCATGCTCTGAGGATGGAGACTGGGAGCGGTGCTGCTGA 618  
 Db 123 ACTCGGGGAGCATTGACCATGCTCTGAGGATGGAGACTGGGAGCGGTGCTGCTGA 182  
 QY 619 AGTCTCAGGCAGAGATATTAACATCCCGAGCGTCCACGTGGGCAAGTCTCCCATGGGTG 678  
 Db 183 AGTCTCAGGCAGAGATATTAACATCCCGAGCGTCCACGTGGGCAAGTCTCCCATGGGTG 242





DT 13-FEB-2002 (first entry)  
 XX DNA encoding novel human diagnostic protein #10554.  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX Homo sapiens.  
 OS WO200175067-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR P-PSDB; ABG10563.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensic, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 10554; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;  
 Query Match 55.0%; Score 651.2; DB 5; Length 2049;  
 Best Local Similarity 92.5%; Pred. No. 3.6e-166;  
 Matches 760; Conservative 0; Mismatches 8; Indels 54; Gaps 5;  
 QY 355 ATTTCCTCGATGATGCTCTCTGAGTCTCTGAGGAAACAATGGGAAGCTCGCCAG 414  
 DB 922 ATTTCCTCGATGATGCTCTCTGAGTCTCTGAGGAAACAATGGGAAGCTCGCCAG 981  
 QY 415 CAGAAGAAATCTCTGCCAAGCCCAAGTTTCAGTTCTCTGTCACAGCCAGGACCTGT 474  
 DB 982 CAGAAGAAATCTCTGCCAAGCCCAAGTTTCAGTTCTCTGTCACAGCCAGGACCTGT 1041  
 QY 475 GACATGAAGCAGAGAGAACAGAGCCACAGCCGTGGCCCTGGCCAGTTTCGGCAGG 534  
 DB 1042 GACCATGAAGCAGAGAGAACAGAGCCACAGCCGTGGCCCTGGCCAGTTTCGGCAGG 1101  
 QY 535 TGGCCCGCCGAGGTGTGCTGTGAGACTCGGGAGCCATTGACCATCTCTCTGAGGATGG 594

DB 1102 TGGCCCGCCGAGGTGTGCTGTGAGACTCGGGAGCCATTGACCATCTCTCTGAGGATGG 1161  
 QY 595 AGACTGGTGGACGGTGTCTGTCTGAAGTCTCAGGCAGAGAGTATAACATCCCCAGCCGTCCA 654  
 DB 1162 AGACTGGTGGACGGTGTCTGTCTGAAGTCTCAGGCAGAGAGTATAACATCCCCAGCCGTCCA 1221  
 QY 655 CGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCCTGAGCAGGAGAGACAGGA 714  
 DB 1222 CGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCCTGAGCAGGAGAGACAGGA 1281  
 QY 715 ACTGTGTTTGTACCTGGGAACCCCTGGAGGGGCTTCTCTCATCCGGAGAGCCAGACCAG 774  
 DB 1282 ACTGTGTTTGTACCTGGGAACCCCTGGAGGGGCTTCTCTCATCCGGAGAGCCAGACCAG 1341  
 QY 775 GAGAGGCTCTTACTCTCTGTCTGAGTCCGGCTCAGCCGCCCTGCATCTCTGGACCGGATCAG 834  
 DB 1342 GAGAGGCTCTTACTCTCTGTCTGAGTCCGGCTCAACCGCCCTGCATCTCTGGACCGGATCAG 1401  
 QY 835 ACACATACAGGATCCACTGCTTGTGACAAATGGCTGGCTGTACATCTCACCCGCGCTCACCTT 894  
 DB 1402 ACACATACAGGATCCACTGCTTGTGACAAATGGCTGGCTGTACATCTCACCCGCGCTCACCTT 1461  
 QY 895 CCCTCTCACTCCAGGCCCTGGTGGACCAATTAATCTGTGAGCTGGCGGATGACATCTGTGCCT 954  
 DB 1462 CCCTCTCACTCCAGGCCCTGGTGGACCAATTAATCTGTGAGCTGGCGGATGACATCTGTGCCT 1495  
 QY 955 ACTCAAGGAGCCCTGTGTCTGTCAGAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCT 1014  
 DB 1496 -----GAGGGCTGGCCGCTCCCTGGCAAGGATATACCCCT 1531  
 QY 1015 ACTGTGACTGTGACAGACACCACTCACTTGGAAAGAGCTGGACAGCTCCCTCTCTGT 1074  
 DB 1532 ACTGTGACTGTGACAGACACCACTCACTTGGAAAGAGCTGGACAGCTCCCTCTCTGT 1591  
 QY 1075 TTCTGAAGCTGCCACA-GGGGAGGAGTCTCTCTCTGAG-TGAGGGTCTCCGGAGT-CCCT 1131  
 DB 1592 TTCTGAAGCTGCCACAGGGGAGGAGTCTCTCTCTGAGGAGGGGCTCCGGAGTCCCT 1651  
 QY 1132 CAGCTTCTACATCAG-CTTGAATGACGAGGCTGTCTCTTTGG 1172  
 DB 1652 CAGCTTCTACATCAGCCCTGTAATGACGAGGCTGTCTCTTTG 1693

## RESULT 10

AAL44087  
 ID AAL44087 standard; cDNA; 1348 bp.  
 XX  
 AC AAL44087;  
 XX  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Mouse modulator of antigen receptor signalling protein coding sequence.

XX Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;  
 KW MARS; tumour suppressor gene; Src-like adaptor protein; SLAP;  
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.

XX Mus sp.

XX Key Location/Qualifiers  
 FH 282..1061  
 FT CDS /tag= a  
 FT /product= "Mouse MARS protein"  
 XX

XX WO200242452-A2.

XX 30-MAY-2002.

XX 26-NOV-2001; 2001WO-CA001662.

XX 27-NOV-2000; 2000CA-02324663.

(HOSP-) HOSPITAL FOR SICK CHILDREN.  
McGlade JC, Loreto MP;  
WPI; 2002-566564/60.  
P-PSDB; AA015456.  
New isolated modulator of antigen receptor signaling protein or its  
fragment, useful for treating malignant disorders such as myeloid  
PT malignancies, autoimmune disorders and myeloproliferative disorders.  
XX  
PS Claim 10; Fig 1A; 11Opp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of modulator  
of antigen receptor signalling (MARS) proteins. The MARS protein is a  
CC putative tumour suppressor gene and exhibits structural and sequence  
CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and  
CC protein sequences of the invention are useful for the treatment of  
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
CC disorders, immunosuppression, myeloproliferative disorders and  
CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
CC breast cancer). The present cDNA sequence encodes a mouse MARS protein  
XX  
SQ Sequence 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;  
  
Query Match 47.5%; Score 561.4; DB 6; Length 1348;  
Best Local Similarity 78.2%; Pred. No. 7.1e-142;  
Matches 714; Conservative 0; Mismatches 191; Indels 8; Gaps 3  
  
QY 274 AGGGCCCCCAAGCCCTAACCTGTGTCAGCCAGACAGCTGCTCTCAGCAGAGCTGTCTTCC 333  
DB 160 AGGACCTGCAAGAGGCTGACCTGTGGGTCAATG--TGACATTTGCTGACTACCTCAT 217  
QY 334 CAAGCCTTTGATGACAAACCAATTTCCCTCGATGATGTGCTTCTGAGTGTCTGTGTAGG 393  
DB 218 CAAACGTCGTATGGCAACCTTTCCCTTTCAGGTTTCAGTGTGCTGTGTGAGGGTCTGTG 277  
QY 394 AACAAATGGGAAGTCTGTGCCAGCAGAGAAGAAATCTCTCCAAAGCCCAAGCTTCGATGTCTCTC 453  
DB 278 AGTGAATGGGAAGTTTGTCCAGCAGAGGGAACCT---CCAGCCCCAGCCCCAGCTCCTC 334  
QY 454 TGTCCAAGGCCAGGAGACCTGTGTGACCATGGAGCAGACAGAGAAGCAGGCCACAGCCGTGGC 513  
DB 335 TGGTCCAGACCCAGGAACCCGGTGTCCATGCATCCAGGAAGACACAAGGTCACAGCTGTGGC 394  
QY 514 CTTGGGCAGTGTTCGGCGCAGGTGGCCCGGCCGAGTGTGCTGTGAGACTCTGGGAGGCCATT 573  
DB 395 CTTGGGCAGTGTTCGCCAGCAGGTGTACAGCCGAGACTATCTGTGAGACTCGGGAGCGCT 454  
QY 574 GACCATCTCTCTGAGGATGGAGACTGTGTGACCGGTGCTGTCTGAAGTCTCAGGCAGAGA 633  
DB 455 GACCATCATCTCTGAGGATGGAGATTGTGTGACAGTCCAGTCGGAAGTCTCAGGCAGAGA 514  
QY 634 GTATAACATCCCCAGCGTCCAGTGGGCAAAAGTCTCCCATGGGTGGTGTGTATGAGGGCT 693  
DB 515 GTACCACATGCCAGTGTGTATGTGGCTAAAGTCGCCACCGGTGGCTGTACAGGGGCT 574  
QY 694 GAGCAGGGAAGAACGACAGGAAGTCTCTGTGTACCTGGGAAACCTTGGAGGGCCCTTCT 753  
DB 575 GAGCCGGGAAGACCGAGGAATCTACTCCTGTATTACCTGGGAACCCCGAGGGGCCCTTCT 634  
QY 754 CATCCGGGAGAGCCAGACAGAGAGGGCTTTACTCTCTGTCACTTCGGCTCAGCCGCC 813  
DB 635 CATCCGGGAGAGCCAGACAGAGAGGGCTGTATTTCCTGTCCGTCCGACTCAGCCGCC 694  
QY 814 TGCATCTGGGACCGGATCAGACTACAGGATCCACTGCTGTTGACATGGCTGGCTGTA 873  
DB 695 TGCATCTGGGACCGGATCAGACACTACAGGATCAGCGCTTTGACAAATGGCTGGCTGTA 754  
QY 874 CATCTCACCGCGCTCACCTTCCCCCTCACTCCAGGCCCTGTGTGACCAATTACTCTGAGCT 933  
DB 755 CATCTCACCTCGGCTCACCTTCCCTCACTCCAGGCTTGGTGGAGCAATTACTCTGAGCT 814



XX DE Human ORF476 coding sequence.  
 XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;  
 XX KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 XX KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 XX KW cancer; cardiovascular disease; allergy; autoimmune disease; ds.  
 XX KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 XX OS Homo sapiens.  
 XX PN US2002082206-A1.  
 XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX PR 30-MAY-2000; 2000US-0208427P.  
 XX PA (LEAC/) LEACH M D.  
 XX PA (MEHR/) MEHRABAN F.  
 XX PA (CONL/) CONLEY P B.  
 XX PA (TOPP/) TOPPER J N.  
 XX PA (LAWD/) LAW D.  
 XX FI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX DR WPI; 2002-626554/67.  
 XX DR P-PSDB; ABP64106.  
 XX DR New polypeptide designated ORFX are present in human atherogenic cells  
 XX PT and are useful to prevent and treat ORFX-associated disorders including  
 XX PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 XX PT inflammatory disease.  
 XX PS Claim 2; SEQ ID NO 951; 78pp; English.  
 XX CC The present invention relates to novel human ORFX polypeptides and their  
 XX CC coding sequences (ABP631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 XX CC were discovered in human atherogenic cells, in particular in platelets  
 XX CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 XX CC many other tissues as well. Atherogenic cells are cells which have the  
 XX CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 XX CC nucleic acids are useful for treating or preventing a pathological  
 XX CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 XX CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 XX CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 XX CC for this patent did not form part of the printed specification, but was  
 XX CC obtained in electronic format directly from the USPTO web site at  
 XX CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
 XX SQ Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;  
 XX CC  
 XX CC Query Match 29.4%; Score 348; DB 6; Length 444;  
 XX CC Best Local Similarity 100.0%; Pred. No. 3.7e-84;  
 XX CC Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 242 CCTAGGACCAAGGACATGCGAGCTTCCAGAGGGCCCAAGCCCTAACCTGTCAG 301  
 DB 1 CCTAGGACCAAGGACATGCGAGCTTCCAGAGGGCCCAAGCCCTAACCTGTCAG 60  
 QY 302 CCAGAGCATGCGTCTCAGCAGAGCTGTCTCCCAAGCCTTTGATGACAAACCAATTCC 361  
 DB 61 CCAGAGCATGCGTCTCAGCAGAGCTGTCTCCCAAGCCTTTGATGACAAACCAATTCC 120  
 QY 362 TCGATGATGCTCTTCTGAGTGCTCTGCTGAGGACATGGAGTGTGCCCACAGAGA 421  
 DB 121 TCGATGATGCTCTTCTGAGTGCTCTGCTGAGGACATGGAGTGTGCCCACAGAGA 180  
 QY 422 AAATCTCTGCCAAGCCCAAGCTTCTCTCTCCAAAGCCAGGACCTGTGACCATG 481  
 DB 181 AAATCTCTGCCAAGCCCAAGCTTCTCTCTCTCCAAAGCCAGGACCTGTGACCATG 240

QY 482 GAAGCAGAGAGAAAGCAAGGCGCAGCCCTGGGCGAGTTTCCGCGAGGTGGCCCG 541  
 DB 241 GAAGCAGAGAGAAAGCAAGGCGCAGCCCTGGGCGAGTTTCCGCGAGGTGGCCCG 300  
 QY 542 GCCGAGCTGTGCTGAGACTCGGGAGGCAATGACCATCGTCTCTGAG 589  
 DB 301 GCCGAGCTGTGCTGAGACTCGGGAGGCAATGACCATCGTCTCTGAG 348  
 RESULT 14  
 ABQ99151  
 ID ABQ99151 standard; DNA; 875 BP.  
 XX AC ABQ99151;  
 XX DT 04-NOV-2002 (first entry)  
 XX DE Human ORF958 coding sequence.  
 XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;  
 XX KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 XX KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 XX KW cancer; cardiovascular disease; allergy; autoimmune disease; ds.  
 XX KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 XX OS Homo sapiens.  
 XX PN US2002082206-A1.  
 XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX PR 30-MAY-2000; 2000US-0208427P.  
 XX PA (LEAC/) LEACH M D.  
 XX PA (MEHR/) MEHRABAN F.  
 XX PA (CONL/) CONLEY P B.  
 XX PA (TOPP/) TOPPER J N.  
 XX PA (LAWD/) LAW D.  
 XX FI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX DR WPI; 2002-626554/67.  
 XX DR P-PSDB; ABP64588.  
 XX PT New polypeptide designated ORFX are present in human atherogenic cells  
 XX PT and are useful to prevent and treat ORFX-associated disorders including  
 XX PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 XX PT inflammatory disease.  
 XX PS Claim 2; SEQ ID NO 1915; 78pp; English.  
 XX CC The present invention relates to novel human ORFX polypeptides and their  
 XX CC coding sequences (ABP631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 XX CC were discovered in human atherogenic cells, in particular in platelets  
 XX CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 XX CC many other tissues as well. Atherogenic cells are cells which have the  
 XX CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 XX CC nucleic acids are useful for treating or preventing a pathological  
 XX CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 XX CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 XX CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 XX CC for this patent did not form part of the printed specification, but was  
 XX CC obtained in electronic format directly from the USPTO web site at  
 XX CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
 XX SQ Sequence 875 BP; 205 A; 259 C; 225 G; 185 T; 0 U; 1 Other;  
 XX CC  
 XX CC Query Match 28.8%; Score 341; DB 6; Length 875;  
 XX CC Best Local Similarity 100.0%; Pred. No. 3.7e-82;  
 XX CC Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:54 ; Search time 7832.77 Seconds  
(without alignments)  
4510.152 Million cell updates/sec

Title: US-09-939-853A-74  
Perfect score: 1183  
Sequence: 1 agctagagctccaaggaccc.....tctcttgatgagcctag 1183

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:

- 1: em\_estba.\*
- 2: em\_estin.\*
- 3: em\_estin.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_htr.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_htr.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pln.\*
- 20: em\_gss\_vrt.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_nam.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_phg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gss1.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	921.4	77.9	1201	9	AL541041
2	860.2	72.7	1020	12	BQ054281
3	817	69.1	1002	12	BQ052308
4	731.4	61.8	1069	12	BQ052468

5	673.6	56.9	778	12	BG178487
6	662.6	56.0	986	12	BQ054265
7	589	49.8	878	12	BQ053486
8	561.4	47.5	2637	11	AK088672
9	546.8	46.2	616	13	BX383606
10	538.6	45.5	2974	11	AK030877
11	476.4	40.3	926	11	AK020837
12	452.2	38.2	660	13	BY742155
13	405	27.3	566	12	BG284179
14	322.4	24.3	537	9	AL844311
15	321	27.1	960	13	BU944126
16	319.8	27.0	660	10	BG635615
17	317.4	26.8	614	9	AL844307
18	314.6	26.6	606	9	AL844309
19	282	23.8	377	9	AA959151
20	278	23.5	627	10	BB619854
c	277.6	23.5	569	14	CB426333
21	272.8	23.1	541	12	BI989899
22	261.4	22.1	794	12	BG677567
23	243.6	20.6	670	9	AI510095
24	243.6	20.2	640	13	BQ553006
25	239.4	19.7	389	10	BE015229
26	233.6	18.7	603	13	BQ553005
c	221.6	18.7	439	10	BE015226
27	221.4	18.7	701	13	BX849096
28	174	14.7	701	13	BX849096
29	169.2	14.3	569	14	CB514688
30	166.8	14.1	642	9	AL844308
c	161.8	13.7	619	9	AL844312
31	161.8	13.7	831	29	AY416277
32	160.8	13.6	871	13	BQ436143
33	159	13.4	570	9	AI471720
34	158.6	13.4	1003	13	BX415149
35	158.6	13.4	1106	13	BX436423
36	158.6	13.4	1133	9	AL549826
37	158.6	13.4	1201	9	AL539427
38	158.6	13.4	1201	9	AL551370
39	158.6	13.4	1083	13	BQ072745
40	157.6	13.3	972	13	BQ707614
41	157.6	13.3	972	13	BQ707614
42	157.4	13.3	849	12	BI769183
43	150.2	12.7	653	13	BY746881
44	149.8	12.7	653	13	BY746881
c	138.4	11.7	611	9	AL844310

## ALIGNMENTS

RESULT 1  
AL541041  
LOCUS AL541041 Homo sapiens PLACENTA Homo sapiens cdna clone CS0DE005YK23  
DEFINITION 5-PRIME, mRNA sequence.  
ACCESSION AL541041  
VERSION AL541041.2 GI:30544829  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
JOURNAL  
COMMENT On Feb 15, 2001 this sequence version replaced gi:12871733.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 131 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9825.r For more information about this cluster, see  
http://www.genoscope.cns.fr/  
csi-bin/cluster.cgi?seq=CS0DE005AF12Q1&cluster=9825.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :

784	QY	TTACTCTCTGTCAGTCGCGCTCAGCGCCCTGCATCCTGGACCGGATCAGACACTACAG	843
845	Db	TTACTCTCTGTCAGTCGCGCTCAGCGCCCTGCATCCTGGACCGGATCAGACCTTAG	904
844	QY	GATCACATGCTTGACAAATGGCTGGCTGTATCATCTCACCGCGCTCAGCTTCCCTCACT	903
905	Db	GATCACATGCTTGACAAATGGCTGGCTGTATCATCTCACCGCGCTCAACTTCCCTCACT	964
904	QY	CCAGCCCTCGTGACCACTTACTTGAGCTGGCGGATGA	942
965	Db	CCAGCCCTCGTGACCACTTACTTGAAGCTGGCGGATGA	1003

```

LOCUS       BQ054281                1020 bp      mRNA      linear      EST 29-MAR-2003
DEFINITION  AGENCOURT_6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:593636
            5', mRNA sequence.
ACCESSION   BQ054281
VERSION     BQ054281.1  GI:19813621
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
            Homo sapiens
REFERENCE   1  (bases 1 to 1020)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
TITLES      Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution by: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCW2125 row: j column: 11
            High quality location sequence stop: 556.
FEATURES             Location/Qualifiers
     1..1020
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:5936362"
         /tissue_type="natural killer cells, cell line"
         /lab_host="DH10B (phage-resistant)"
         /clone_lib="NIH MGC 106"
         /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

```

ORIGIN		nir_106c library.	
Query	Katch	72.7%;	Score 860.2; DB 12; Length 1020;
Best Local Similarity	96.1%;	Pred. No. 3.9e+206;	
Matches	956;	Conservative 0;	Mismatches 30; Indels 9; Gaps 7
Qy	187	ACCAACACTAGCCCTCCCTGAGATCCTCCAGGCTGAGAGAGTTCTGGTGTCCTAG	246
Db	1	AACCAACCTAGCCCTCCCTGAGATCCTCCAGGCTGAGAGTTCTGGTGTCCTAG	60
Qy	247	GACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTCTCCAGCCAGA	308
Db	61	GACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAACCTGTCTCCAGCCAGA	120
Qy	307	GCATCGCTCTCAGCAGAGCTGCTTCTCCCAAGCCCTTGATGACAAACCAANTTCCCTCGAT	366

Db 121 GCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCGAT 180  
Qy 367 CATGTGCTTCTGAGTCTCTCTCAGGAAACAATGGGAAGTCTGCCAGCAGAGAAATC 426  
Db 181 GATGTGCTTCTGAGTCTCTCTCAGGAAACAATGGGAAGTCTGCCAGCAGAGAAATC 240  
Qy 427 TCTGCCAAGCCCAAGCTTGAATTCCTCTGTCTCAAGGCCAGGACCTGTGACCATGGAAGC 486  
Db 241 TCTGCCAAGCCCAAGCTTGAATTCCTCTGTCTCAAGGCCAGGACCTGTGACCATGGAAGC 300  
Qy 487 AGAGAGAAGCAAGGCCACAGCGTGGCCCTGGGCGAGTTTCCCGCAGCTGGCCGCGCA 546  
Db 301 AGAGAGAAGCAAGGCCACAGCGTGGCCCTGGGCGAGTTTCCCGCAGCTGGCCGCGCA 360  
Qy 547 GCTGTGCTGAGATCTGGGGAGCATTGACCATCTCTGAGGATGAGACCTGTTGAC 606  
Db 361 GCTGTGCTGAGATCTGGGGAGCATTGACCATCTCTGAGGATGAGACCTGTTGAC 420  
Qy 607 GGTGCTGTCAAGTCTCAGCAGAGAGTATAACATCCCGAGCTGCCAGTGGGCAAGT 666  
Db 421 GGTGCTGTCAAGTCTCAGCAGAGAGTATAACATCCCGAGCTGCCAGTGGGCAAGT 480  
Qy 667 CTCCATGGTGGTGTATGAGGCGCTGAGCAGGAGAAAGCAGAGAACTGCTGTTGTT 726  
Db 481 CTCCATGGTGGTGTATGAGGCGCTGAGCAGGAGAAAGCAGAGAACTGCTGTTGTT 540  
Qy 727 ACCTGGGAACCTCGAGGGGCTTCTCATCCGGAGAGCCAGACAGGAGAGCTCTTA 786  
Db 541 ACCTGGGAACCTCGAGGGGCTTCTCATCCGGAGAGCCAGACAGGAGAGCTCTTA 600  
Qy 787 CTCTCTGTGAGTCCGCTCAGCCGCTGATCTGGGACCGGATCAGACACTACAGGAT 846  
Db 601 CTCTCTGTGAGTCCGCTCAGCCGCTGATCTGGGACCGGATCAGACACTACAGGAT 660  
Qy 847 CCATGCTTGACAAATGGCTGGTGTACATCTCACCGGCTCACCCTTCCCTCACTCCA 906  
Db 661 CCATGCTTGACAAATGGCTGGTGTACATCTCACCGGCTCACCCTTCCCTCACTCCA 720  
Qy 907 GG-CCCTGGTGGACCAATTACTCTGAGTGGGGGATGACATCTGTGCTA-CTCAAGGAG 964  
Db 721 GGCCCTGGTGGACCAATTACTCTGAGTGGGGGATGACATCTGTGCTANTCTCAGGAC 780  
Qy 965 CCCTGTGTCTGAGAGGGCTGGCCGCTCCTCGCAAGGATATA-CCCTTACTGTGAC 1023  
Db 781 CCCTGTGTCTGCAAGGGCTGGCCGCTCCTCGCAAGGATATA-CCCTTACTGTGAC 840  
Qy 1024 TGTGAGAGGACCACTCACTGAGAGAGAGTGGAGAGTGGAGAGTGGAGTGGAGTGGAG 1081  
Db 841 TGTGAAAGGACCCACTCACTGGGAGAGTGGAGAGTGGAGTGGAGTGGAGTGGAG 900  
Qy 1082 GCTGCCACA-GGGGAGAGTCTCTTCTCAGTGAGG--TCTCCGGAGTCCCTCAGCTTC 1138  
Db 901 GCTGCCCCAGGGGAGAGTCTCTTCTCATTGAGGGGTCTCCGGAGTCCCTCAGCTTC 960  
Qy 1139 TACATCAGCTG-AATGACGAGGTGTCTCTTTGG 1172  
Db 961 TACATCAACCTGTATGACCAAGGCTGGCTTTG 995

RESULT 3  
LOCUS BQ052308 1002 bp mRNA linear EST 29-MAR-2002  
DEFINITION AGENCOURT\_686571 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5933542  
5', mRNA sequence.  
ACCESSION BQ052308  
VERSION BQ052308.1 GI:19811648  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1002)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
cDNA Library Preparation: Ruben Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2118 row: d column: 23  
High quality sequence stop: 670.

FEATURES  
source  
1..1002  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5933542"  
/tissue\_type="natural killer cells, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 106"  
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

ORIGIN  
Query Match 69.1%; Score 817; DB 12; Length 1002;  
Best Local Similarity 98.3%; Fred. No. 3.2e-195;  
Matches 868; Conservative 0; Mismatches 10; Indels 5; Gaps 4;  
Qy 302 CCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCAGCGCTTTGATGACAAACCAATTTCCC 361  
Db 1 CCAGAGCATGCGTCTCAGCAGAGCTGTCTT-CCAAGCCTTTGATGACAAACCAATTTCCC 59  
Qy 362 TCGATGATGTGCTTCTGAGTGTCTGTGAGGAACAATGGGAAGTCTGCCAGCAGAGA 421  
Db 60 TCGATGATGTGCTTCTGAGTGTCTGTGAGGAACAATGGGAAGTCTGCCAGCAGAGA 119  
Qy 422 AAATCTCTGCCAAGCCCAAGCTTGAGTTTCTCTGTCGAAGCCAGGACCTGTGACCATG 481  
Db 120 AAATCTCTGCCAAGCCCAAGCTTGAGTTTCTCTGTCAGGCCAGGACCTGTGACCATG 179  
Qy 482 GAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTGGGCGAGTTTCCCGGAGGTGGCCCG 541  
Db 180 GAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTGGGCGAGTTTCCCGGAGGTGGCCCG 239  
Qy 542 GCCGAGCTGTGCTCAGACTCGGGAGGCCATTGACCATGCTCTGTAGGATGGAGACTGG 601  
Db 240 GCCGAGCTGTGCTCAGACTCGGGAGGCCATTGACCATGCTCTGTAGGATGGAGACTGG 299  
Qy 602 TGGACGGTGTGTCTGAAAGTCTCAGCAGAGAGTATACATCCCGAGCTCCAGCTGGGCG 661  
Db 300 TGGACGGTGTGTCTGAAAGTCTCAGCAGAGAGTATACATCCCGAGCTCCAGCTGGGCG 359  
Qy 662 AAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGCGGAGAGAGCAGAGAACTGCTG 721  
Db 360 AAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGCGGAGAGAGAGAGAACTGCTG 419  
Qy 722 TTGTTTACCTGGGAACCTTGGAGGGCCCTTCTCTCATCCGGGAGAGCCAGACAGAGAGGC 781  
Db 420 TTGTTTACCTGGGAACCTTGGAGGGCCCTTCTCTCATCCGGGAGAGCCAGACAGAGAGGC 479  
Qy 782 TCTTACTCTCTGCTGAGTCCGCTCAGCGCGCTGATCTCTGGACCGGATCAGACACTAC 841  
Db 480 TCTTACTCTCTGCTGAGTCCGCTCAGCGCGCTGATCTCTGGACCGGATCAGACACTAC 539  
Qy 842 AGGATCCACTGCTCTTGAACAATGGGTGGCTGTATCTCTCAGCGCGCTCACCTTCCCTCA 901

```

Db 540 AGGATCTACTGCTTGGACAAATGGCTGGCTGATCATCTACCGGCGCTACCTTCCCTCA 599
Qy 902 CTCAGGCCCTGGTGACCAATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTCAAG 961
Db 600 CTCAGGCCCTGGTGACCAATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTCAAG 659
Qy 962 GAGCCCTGTGCTGTCGAGAGGGTGCGCGCTCCCTGCGCAAGGATATACCCCTACTCTG 1021
Db 660 GAGCCCTGTGCTGTCGAGAGGGTGCGCGCTCCCTGCGCAAGGATATACCCCTACTCTG 719
Qy 1022 ACTGTGAGAGGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAA 1081
Db 720 ACTGTGAGAGGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAA 779
Qy 1082 GCTGCCACA-GGGAGGAGTCTTCTCAGTAGGG--TCTCGGGAGTCCCTCAGCTTC 1138
Db 780 ACTGCCACAGGGGAGGAGTCTTCTCAGTAGGGGTCTCCGGGAGTCCCTCAGCTTC 839
Qy 1139 TACATCAGCTCGAATGACGA-GGCTGTCTCTTTGGATGATGCC 1180
Db 840 TACATCAGCTCGAATGACGA-GGCTGTCTCTTTGGATGATGCC 882

RESULT 4
BQ052468 1069 bp mRNA linear EST 29-MAR-2002
LOCUS AGENCOURT 6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
5', mRNA sequence.
ACCESSION BQ052468
VERSION BQ052468.1 GI:19811808
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1069)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rsapbs@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2118 row: n column: 13
High quality sequence stop: 681.
Location/Qualifiers
1. 1069
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5933772"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site:1; XhoI; Site:2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

Query Match 61.8%; Score 731.4; DB 12; Length 1069;
Best Local Similarity 94.0%; Pred No. 1.4e-173;
Matches 824; Conservative 0; Mismatches 1; Indels 52; Gaps 4;

ORIGIN

```

```

Qy 310 TGGCTCTCAGCAGAGCTGTCTTCCCAAGCCCTTTGATGACAAACCAATTTCCCTCGATGAT 369
Db 1 TGGCTCTCAGCAGAGCTGTCTTCCCAAGCCCTTTGATGACAAACCAATTTCCCTCGATGAT 60
Qy 370 GTGCTTCTGAGTGTCTGCTGAGGACAAATGGGAAGTCTCCCGAGCAGAGAGAAATCTCT 429
Db 61 GTGCTTCTGAGTGTCTGCTGAGGACAAATGGGAAGTCTGCGCCAGCAGAGAAATCTCT 120
Qy 430 GCCAAGCCCAAGCTTGAGTTCCTCTGTCCAAAGCCAGGAGACCTGTGACCATGAAGCAGA 489
Db 121 GCCAAGCCCAAGCTTGAGTTCCTCTGTCCAAAGCCAGGAGACCTGTGACCATGAAGCAGA 180
Qy 490 GAGAAGCAAGCCACACGCGCTGGCGGCTGGCGGAGTTCCTCCCGGAGGTGGCCCGCGAGCT 549
Db 181 GAGAAGCAAGCCACACGCGCTGGCGGCTGGCGGAGTTCCTCCCGGAGGTGGCCCGCGAGCT 240
Qy 550 GTGCTGAGACTGTGGGAGCCATTGACCATCTGCTCTGAGGATGGAGACTGTGGTGGACGGT 609
Db 241 GTGCTGAGACTGTGGGAGCCATTGACCATCTGCTCTGAGGATGGAGACTGTGGTGGACGGT 300
Qy 610 GCTGCTGAAAGTCTCAGCAGAGAGATATAACATCCCGCAGGTCCACGTGGGCAAGTCTC 669
Db 301 GCTGCTGAAAGTCTCAGCAGAGAGATATAACATCCCGCAGGTCCACGTGGGCAAGTCTC 360
Qy 670 CCATGGGTGGCTGTATGAGGGCTGTGACGAGGAGAAAGCAGAGGAACTGTGTGTGTACC 729
Db 361 CCATGGGTGGCTGTATGAGGGCTGTGACGAGGAGAAAGCAGAGGAACTGTGTGTGTACC 420
Qy 730 TGGGAACCCCTGGAGGGGCTTCTCATCCGGGAGAGCCAGCAGGAGAGGCTCTTACTC 789
Db 421 TGGGAACCCCTGGAGGGGCTTCTCATCCGGGAGAGCCAGCAGGAGAGGCTCTTACTC 480
Qy 790 TCTGTGAGTCCGCTCAGCGGCTTGTGATCTGGGAGCTGATGATCTGCTGCTACTCAAGGAGCCCT 968
Db 601 CCTGGTGGACCAATTAATCTGAG----- 623
Qy 969 GTGCTCTGACAGAGGCTGGCGCGCTCCCTGCGCAAGGATATACCCCTACTCTGTGACTGTC 1028
Db 624 -----GGGCTGGCGCGCTCCCTGCGCAAGGATATACCCCTACTCTGTGACTGTC 671
Qy 1029 AGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAAGCTGCCA 1088
Db 672 AGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAAGCTGCCA 731
Qy 1089 CA-GGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGG-AGTCCCTCAGTCTCTACATCAG 1146
Db 732 CAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGTCTCTACATCAG 791
Qy 1147 CCTGAATGACGAGGCTGTCTCTTTGGATGATGCCCTAG 1183
Db 792 CCTGAATGACGAGGCTGTCTCTTTGGATGATGCCCTAG 828

```

```

RESULT 5
BQ178487 778 bp mRNA linear EST 06-FEB-2001
LOCUS 602328305F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429896 5',
DEFINITION mRNA sequence.
ACCESSION BQ178487
VERSION BQ178487.1 GI:12685190
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE  
 AUTHORS NIH-MGC http://img.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: DCTB/DTF  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM10182 row: i column: 01  
 High quality sequence stop: 657.  
 Location/Qualifiers  
 1..778  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4429896"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_91"  
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."  
 ORIGIN  
 Query Match 56.9%; Score 673.6; DB 12; Length 778;  
 Best Local Similarity 96.8%; Pred. No. 4.8e-159;  
 Matches 730; Conservative 0; Mismatches 19; Indels 5; Gaps 4;  
 QY 36 GTGACAGCTCAAGGGCCCTGGSCCTTCCTCCCTGGCTGGCTGGCTGGAGGGT 95  
 Db |||||  
 QY 1 GTGACAGCTCAAGGGCCCTGGG-CTTCCTCCCTGGCTGGCTGGAGGGT 59  
 Db |||||  
 QY 96 TCCCGAGTCCGAATCCCTTAAGGAGCATGGGCGAGCTGATCCATCCCTGGTGTACAACT 155  
 Db |||||  
 QY 60 TCCCGAGTCCGAATCCCTTAAGGAGATGGGCGAGCTGATCCATCCCTGGTGTACAACT 119  
 Db |||||  
 QY 156 GTGACTGACAGAGATGCTAGCTACCCAAACCAACCTAGCTCTCCCTGGAGATCC 215  
 Db |||||  
 QY 120 GTGACTGACAGAGATGCTAGCTACCCAAACCAACCTAGCTCTCCCTGGAGATCC 179  
 Db |||||  
 QY 216 TCCCGAGTCCGAATCCCTTAAGGAGCATGGGCGAGCTGATCCATCCCTGGTGTACAACT 275  
 Db |||||  
 QY 180 TCCCGAGTCCGAATCCCTTAAGGAGATGGGCGAGCTGATCCATCCCTGGTGTACAACT 239  
 Db |||||  
 QY 276 GGGCCCCAAGCCCTAACCTGTCAGCCAGAGCATGGCTCTCAGCAGAGCTGTCTTCCCA 335  
 Db |||||  
 QY 240 GGGCCCCAAGCCCTAACCTGTCAGCCAGAGCATGGCTCTCAGCAGAGCTGTCTTCCCA 299  
 Db |||||  
 QY 336 AGCTTTGATGACCAACCAATTTCCCTCGATGATGCTTCTCGAGTGTCTCGCTGAGGAA 395  
 Db |||||  
 QY 300 AGCTTTGATGACCAACCAATTTCCCTCGATGATGCTTCTCGAGTGTCTCGCTGAGGAA 359  
 Db |||||  
 QY 396 CAATGGCAAGTGTGCCAGCAGAGAAATCTCTCCAGCCCAAGCTTGAGTTCTCTCTG 455  
 Db |||||  
 QY 360 CAATGGCAAGTGTGCCAGCAGAGAAATCTCTCCAGCCCAAGCTTGAGTTCTCTCTG 419  
 Db |||||  
 QY 456 TCCAGGCCAGGAGCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 515  
 Db |||||  
 QY 420 TCCAGGCCA-GGACCTGTGACCATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478  
 Db |||||  
 QY 516 TGGGCAAGTTTCCCGGAGGTGCCCGCGGCGAGCTGTCTCGTGTGAGACTCGGGGAGCCATTGA 575  
 Db |||||  
 QY 479 TGGGCAAGTTTCCCGGAGGTGCCCGCGGCGAGCTGTCTCGTGTGAGACTCGGGGAGCCATTGA 538  
 Db |||||  
 QY 576 CCATCGTCTCTGAGGATGAGACTGGTGTGACCGGTGCTGTCTGAAAGTCTCAGGCAGAGAGT 635  
 Db |||||

Db 539 CCATCGTCTCTGAGGATGGAGACTGGTGGCGGTGCTGTGAAGTCTCAGGCAGAGT 598  
 QY 636 ATAACATCCCCAGCGTCCACGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGA 695  
 Db 599 ATAACATCCCCAGCGTCCACGTGGGCAAGTCTCACAATGGTGGCTGTATGAGGGCTGA 658  
 QY 696 GCAGGGAGA-AAGCAGAGGAAGTCTGTGTACCTGGGAAACCTCGAGGGGCTTCC-- 752  
 Db 659 GCAGGGAGACAAGCAGAGGAAGTCTGTGTACCTGGGAAACCTCGAGGGGCTTCCCT 718  
 QY 753 TCATCCGGGAGAGCCAGACAGCAGAGAGGCTCTTA 786  
 Db 719 CATCCCGAAGAGCAGACAGCAGAGAGGCTCTTA 752  
 RESULT 6  
 BQ054265  
 LOCUS AGENCOURT 6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BQ054265  
 VERSION BQ054265.1 GI:19813605  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 986)  
 NIH-MGC http://img.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM2125 row: i column: 12  
 High quality sequence stop: 515.  
 Location/Qualifiers  
 1..986  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5936339"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_106"  
 /notes="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACACG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
 ORIGIN  
 Query Match 56.0%; Score 662.6; DB 12; Length 986;  
 Best Local Similarity 98.6%; Pred. No. 3.4e-156;  
 Matches 721; Conservative 0; Mismatches 4; Indels 6; Gaps 5;  
 QY 120 GCATGGGCGAGCTGATCCATCCCTGGTGTACAACTGTGCTGACAGAGATGCTGAGC 179  
 Db |||||  
 QY 1 GCATGGGCGAGCTGATCCATCCCTGGTGTACAACTGTGCTGACAGAGATGCTGAGC 60  
 Db |||||  
 QY 180 TACCCAAACCAACCACTAGCTCTCCCTGAAGATCTCCAGGCTCAGAGAGTCTGGGT 239  
 Db |||||  
 QY 61 TACCCAAACCAACCACTAGCTCTCCCTGAAGATCTCCAGGCTCAGAGAGTCTGGGT 120  
 Db |||||

QY	240	GTCTAGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTAACCTGTGC	299
Db	121	GTCTAGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTAACCTGTGC	180
QY	300	AGCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTTC	359
Db	181	AGCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTTC	240
QY	360	CCTGATGATGTGTTCTTGATGCTCTGCTGAGGAAACATGGAAGTCTGCCAGAGAA	419
Db	241	CCTGATGATGTGTTCTTGATGCTCTGCTGAGGAAACATGGAAGTCTGCCAGAGAA	300
QY	420	GAAATCTTGCCAAAGCCCAAGCTTGAGTTCTCTGTCACAGCCAGGACCTGTGACCA	479
Db	301	GAAATCTTGCCAAAGCCCAAGCTTGAGTTCTCTGTCACAGCCAGGACCTGTGACCA	360
QY	480	TGGAAGCAGAGAGAAAGCAAGCCACAGCCGTGGCCCTGGGCACTTTCCCGCAGGTGGCC	539
Db	361	TGGAAGCAGAGAGAAAGCAAGCCACAGCCGTGGCCCTGGGCACTTTCCCGCAGGTGGCC	420
QY	540	CGGCCGAGCTGTGCTCAGACTCGGGAGCCATTGACCATGCTCTGAGGATGGAGACT	599
Db	421	CGGCCGAGCTGTGCTCAGACTCGGGAGCCATTGACCATGCTCTGAGGATGGAGACT	480
QY	600	GGTGGACGGTCTGCTCTGAAAGTCTCAGGACAGAGATATAACATCCCGACGCTCCACGTGG	659
Db	481	GGTGGACGGTCTGCTCTGAAAGTCTCAGGACAGAGATATAACATCCCGACGCTCCACGTGG	540
QY	660	GCAAGTCTCCATGGGTGGCTGTATCAGAGGCTGATCAGGAGAGAGAGAGGAACTGC	719
Db	541	CCAAGTCTCCATGGGTGGCTGTATCAGAGGCTGATCAGGAGAGAGAGAGGAACTGC	600
QY	720	TGTTGTTTACCTGGGAACCCCTGGAGGGG- CCTTCTCTCATCCGGAG- AGCCAGACCCAGGAG	777
Db	601	TGTTGTTTACCTGGGAACCCCTGGAGGGGCTTCTCTCATCCGGAGAGCCAGACCCAGGAG	660
QY	778	AGGCTCTTACTCTCTGTCAGT- CGGCTCTAGCG- CCTGTCATCTCTGGG- ACCGATCA	833
Db	661	AGGCTCTTACTCTCTGTCAGTCTCGGCTCTGAGCGCCCTGATCTCTCTGGGACCGGATCA	720
QY	834	GACACTACAGG 844	
Db	721	GACCTTACAGG 731	
RESULT 7			
BQ053486			
LOCUS			
DEFINITION BQ053486 6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253			
5', mRNA sequence.			
ACCESSION BQ053486			
VERSION BQ053486.1 GI:19812826			
KEYWORDS EST.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 878)			
NIH-MGC <a href="http://mgs.nci.nih.gov/">http://mgs.nci.nih.gov/</a> .			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a>			
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI			
CDNA Library Preparation: Rubin Laboratory			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
<a href="http://image.llnl.gov">http://image.llnl.gov</a>			
Plate: LLCM2122 row: 1 column: 06			
High quality sequence stop: 394.			

FEATURES		Location/Qualifiers	
source		1..878	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5935253"	
		/tissue_type="natural killer cells, cell line"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_106"	
		/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN		Query Match 49.8%; Score 589; DB 12; Length 878;	
		Best Local Similarity 95.4%; Pred. No. 1.2e-137;	
		Matches 660; Conservative 0; Mismatches 26; Indels 6; Gaps 5;	
QY	125	GGGAGCTGATCCATCCCTGCTGATACAACTGCTGACGAGATGCTGACGATGCTGACTACCC	184
Db	1	GGGAGCTGATCCATCCCTGCTGTA-AAACTGCTGACTGCAGACAGATGCTGATCTACCC	59
QY	185	AAACCAACACCTAGCTCTCCCTGAAGATCCTCCAGGCTGAGAGAGTCTGGGTGTCTT	244
Db	60	AAACCAACACCTAGCTCTCCCTGAAGATCCTCCAGGCTGAGAGAGTCTGGGTGTCTT	119
QY	245	AGGACCAAGACACTGGCAGACTTCCAGAGGGCCCCAAAGCCCTAACCTGTCCAGCCA	304
Db	120	AGGACCAAGACACTGGCAGACTTCCAGAGGGCCCCAAAGCCCTAACCTGTCCAGCCA	179
QY	305	GAGCATGGTCTCAGCAGAGCTGCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCG	364
Db	180	GAGCATGGTCTCAGCAGAGCTGCTTCCCAAGCCTTTGATGACAAACCAATTTCCCTCG	239
QY	365	ATGATGTCCTTCTGAGTGTCTCTGAGGAACAAATGGGAAGTCTGCCAGCAGAGAAA	424
Db	240	ATGATGTCCTTCTGAGTGTCTCTGAGGAACAAATGGGAAGTCTGCCAGCAGAGAAA	299
QY	425	TCTTGCCCAAGCCCAAGCTTGAGTTCCTCTGTCAAGCCAGGACCTGTGACCATGAA	484
Db	300	TCTTGCCCAAGCCCAAGCTTGAGTTCCTCTGTCAAGCCAGGACCTGTGACCATGAA	359
QY	485	GCAGAGAAAGCAAGGCCACAGCCGTGGCCCTGGCGAGTTTCCCGCAGGTGGCCCGGCC	544
Db	360	GCAGAGAAAGCAAGGCCACAGCCGTGGCCCTGGCGAGTTTCCCGCAGGTGGCCCGGCC	419
QY	545	GAGCTGTGCTGAGACTCGGGAGCCATTGACCATGCTCTGAGGATGGAGCTGTGG	604
Db	420	GAGCTGTGCTGAGACTCGGGAGCCATTGACCATGCTCTGAGGATGGAGCTGTGG	479
QY	605	ACGCTGTGCTGAAAGTCTCAGGCAGAGATATACATCCCGCAGCTCCAGTGGGCAAA	664
Db	480	ACGCTGTGCTGAAATCTCAGGCAGAGATATACATCCCGCAGCTCCAGTGGGCAAA	539
QY	665	GTCTCCCATGGTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGACTCTGTTG	724
Db	540	GTCTCCCATGGTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGACTCTGTTG	597
QY	725	TTACTCTGGAAACCTTGGAGGGCCCTTCTCAT-CCGGGAGAGCCAGACGAGAGGCT-	782
Db	598	TTACTCTGGAAACCTTGGAGGGCCCTTCTCAT-CCGGGAGAGACACAGACGAGAGGCTC	657
QY	783	CTTACTCTCTGT-CAGTCCGCTCAGCGGCC 813	
Db	658	CTAATCTCTGTGTAGTCCGCTTTCAGCGGCC 689	
RESULT 8			

AK088672 2637 bp mRNA linear HTC 20-SEP-2003  
 LOCUS Mus musculus 2 days neonate thymus cells cDNA, RIKEN  
 DEFINITION full-length enriched library, clone:E430023D24 product:MODULATOR OF  
 ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.  
 ACCESSION AK088672.1 GI:26353729  
 VERSION  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carninci, P. and Hayashizaki, Y.  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159  
 REFERENCE  
 2  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multipipillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 REFERENCE  
 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 695-690 (2001)  
 REFERENCE  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 REFERENCE  
 6 (bases 1 to 2637)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Cheato, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216]  
 COMMENT cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
 Trust/MRC building Addenbrookes Hospital Cambridge) whose  
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
 assistance we gratefully acknowledge.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/.

FEATURES  
source

1 2637

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="NOD"

/db\_xref="FANTOM DB:E430023D24"

/db\_xref="MGI:2427569"

/db\_xref="taxon:10090"

/clone="B430023D24"

/cell\_type="thymic cells"

/tissue\_type="thymus"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="2 days neonate"

358..1137

/note="unnamed protein product; MODULATOR OF ANTIGEN

RECEPTOR SIGNALING MARS (SPTP)[AAL38196, evidence: FASTA,

100%ID, 100%length, match=777]

putative"

/codon\_start=1

/protein\_id="BAC40495.1"

/db\_xref="GI:26353730"

/translation="MGSLSRGKTSPPSSGPDQVPMOPERHKVTAVALGSPPA

GEARLSALRLGPLELTIISDDGWTWVQSVSGREVHPMSVYAKVAHGLVGLSREK

AEELLLPCNPGSAFLIRESCYSLVSLRSPASWDRIHVPLOLDNGLYIS

PLRTFSLALVEHVSLEADGCCPLREPCVLQKLGPLPKDTPPVTPTVTSLNWK

LDRSFLLEAPASGEASLSEGLRESLSYSIASEDPLDDA"

## ORIGIN

Query Match 47.5%; Score 561.4; DB 11; Length 2637;

Best Local Similarity 78.2%; Pred. No. 2.4e-130;

Matches 714; Conservative 0; Mismatches 191; Indels 8; Gaps 3;

QY 274 AGGCCCCCAAGCCCTAACCTGTCAGCAGAGCATGCGTCTCAGCAGAGCTGTCTTC 333

DB 236 AGGACCTGCAAGGCGCTGACCTGTGGGTCACTG--TGCACATTTGCTGACTACCTCAT 293

QY 334 CAAGCCTTTGATGACAAACCAATTTCCCTCGATGATGCTCTCTGAGTCTCTGCTGAGG 393

DB 294 CAACGCTGATGGCAACCTTTCCCTTCCAGTTCACTGCTGCTGAGCGTCTGCTG 353

QY 394 AACATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTC 453

DB 354 AGTGATGGGAAGTTTGTCCAGCAGAGAGGAAAACCT---CCAGCCCCAGCCAGCTCCTC 410

QY 454 TGTCGAAGCCAGGACCTGTGACCATGGAACAGAGAGAACGACAGCCACAGCCGTGGC 513

DB 411 TGTCCAGACACAGAACCCGCTGTCATGCAACCCAGAAAGACACAGCTCACAGCTGTGGC 470

QY 514 CTGGGCGAGTTTCCCGGCGAGGTGGCCCGCCGAGTGTGCTGCTGAGACTCGGGAGCAATT 573

DB 471 CTGGGCGAGTTTCCCGGCGAGGTGAAACAGCCAGACTATCTCTGAGACTCGGGAGCCGCT 530

QY 574 GACCATCGTCTCTGAGGATGGAGACTGTGTGGAGCGGTGTGTCTGAAGTCTCAGGAGAGA 633

DB 531 GACCATCATCTCTGAGGATGGAGATTTGGTGGACAGTCCAGTCCGAGTCTCAGGAGAGA 590

QY 634 GTATAACATCCCGAGCGTCCAGCTGGGCAAGTCTCCATGCTGGTGGCTGTATGAGGGCCT 693

DB 591 GTACCACATGCCAGTGTGTATGTGCTAAAGTCCGCCACGCGGTGCTGTACAGGGCCT 650

QY 694 GAGCAGGAGAGAACAGAGAGAACTGCTGTGTATCTGGGAAACCTCTGAGGGGCGCTTCT 753

DB 651 GAGCCGGGAGAGAACCGAGGAACTACTCTGTGTACCTGGGAAACCCCGAGGGGCGCTTCT 710



prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
2049374  
11042159

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5

# TITLE

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 2974)

# JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Ahira, S.,  
Takeda, Y., Tanaka, I., Tomaru, A., Toyota, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

# TITLE

# JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan [E-mail: genome-res@gs.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216]

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

## FEATURES

### source

1. 2974  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:5830437K10"  
/db\_xref="MGI:2394046"  
/db\_xref="taxon:10090"  
/clone="5830437K10"  
/sex="male"  
/tissue\_type="thymus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stages="adult"  
232. .1011  
/note="unnamed protein product; MODULATOR OF ANTIGEN  
RECEPTOR SIGNALING MARS (SPTR)AAL38196, evidence: PASTY,

## CDS

100%ID, 100%length, match=777)  
putative"  
/codon\_start=1  
/protein\_id="BAC27168.1"  
/db\_xref="GI:26326949"  
/translation="MGSLSRSGTKSSPSSSPDQEPVSMOPERHKVTAVALGSPPA  
GEARLSRLGELPTIISDGDWMTVQSVSGREVHMPVYVAKVAHGLYGLSSEK  
AEELLILPFGGAFILRSQTRRGYSLSVRLSPASWDRIRHRIQLRDNGWLYIT  
PRUTFSLHALVEHSELADGICCPUREPCVQKUGLPKDTPPPTVPTVTSLSNWK  
LDRSLFLFAPASGEASLLSEGLRSLSSYIISLAEDPLDDA"

## ORIGIN

Query Match 45.5%; Score 538.6; DB 11; Length 2974;  
Best Local Similarity 80.9%; Pred. No. 1.5e-124;  
Matches 653; Conservative 0; Mismatches 148; Indels 6; Gaps 2;  
QY 380 GTGCTCTGTCAGAAACAATGGGAATCTGCCAGCAGAGAATAAATCTCTGCCAACCCCA 439  
DB 214 GTGAGCGTCTGCTGAGTGATGGGAATTTGTCAGCAGAGGGAAC---CTCCAGCCCC 270  
QY 440 AGTTTCAGTTCTCTCTCAAGGCCAGGACCTGTGACCATGGAACGAGAGAGAACAG 499  
DB 271 AGCCCCAGTCTCTCTGTCAGACAGAACCGTGTCCATGCCAACCAAGAACACAG 330  
QY 500 GCACAGCCGTGGCCCTGGGCAGTTTCCCGCAGAGTGCCCGCCGAGTGTGCTGAGA 559  
DB 331 GTCACAGCTGTGGCCCTGGGCAGTTTCCCGCAGAGTGTAACAGGCCAGACTATCTCTG 390  
QY 560 CTCGGGGGCCATTGACCATCTCTCTGAGGATGGAGTGGTGGAGGCTGTCTGAA 619  
DB 391 CTCGGGGAGCGGTGACCATCATCTCTGAGGATGGAGATTTGGGACATTCAGTCCGAA 450  
QY 620 GTCTCAGGAGAGATTAACATCCCGACGCTCCAGTGGGAAAAGTCTCCATGGGTGG 679  
DB 451 GTCTCAGGAGAGATTAACATCCCGACGCTCTGATGTGCTAAAGTCCGCCACGGGTGG 510  
QY 680 CTGTATGAGGGCTGACGAGGAGAGAAAGCAGAGGAACCTCTGTTTACTTGGGAACCT 739  
DB 511 CTGTACGAGGGCTGAGCCGGGAGAAACCGGAGGAACCTACTCTGTACTGGGAACCCC 570  
QY 740 GGAGGGGCTTCTCTCATCCGGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTGAGTC 799  
DB 571 GGAGGGGCTTCTCTCATCCGGGAGAGCCAGACAGGAGAGGCTACTATTCCTGTCCGTC 630  
QY 800 CGCTCAGCGGCTGCTCATCTCTGGGACCGGATCAGACTACAGGATCCACTGCTTGAC 859  
DB 631 CGACTCAGCGGCTGCTCATCTCTGGGACCGGATCAGACTACAGGATCAGCGTCTTGAC 690  
QY 860 AATGGTGGCTGTACATCTCACCGGCTCACTTCCCTCACTCCAGGCCCTGTGGAC 919  
DB 691 AATGGTGGCTGTACATCAGCTGCTCACTTCCCTCACTCCAGGCCCTGTGGAG 750  
QY 920 CATTAATCTGAGCTGGCGGATGACATCTGCTGCTCACTCAAGAGCCCTGTGCTGCGAG 979  
DB 751 CATTAATCTGAGCTGAGATGGCATCTGCTGCTCCCTCAGGAGCCCTGTGCTGCTGAG 810  
QY 980 AGGGCTGGCCGCTCCCTGGGCAAGGATATACCCCTACCTGTGACCTGTCAGAGGACACA 1039  
DB 811 AGCTTGGGCCACTACCTGGGCAAGATACCTCCACTGTGACTGTGCCAATCATCA 870  
QY 1040 CTCAACTGGAAAGAGCTGGACAGCTCCCTCTGTTTCTGAG---CTGCCACAGGGAG 1096  
DB 871 CTAAATTGGAAAAGAGCTGGACCGCAGGCTCTCTGTTTCTGGAAGACCTGCGAGTGGGAG 930  
QY 1097 GAGTCTCTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGCTGATGAC 1156  
DB 931 GCACTCTCTCAGTGAGGGGCTCCGAGAGTCCCTCAGTCTCTCATCAGCTGCTGAG 990  
1157 GAGGCTGTCTCTTTTGATGATGCTTAG 1183  
991 GACCCCTTGGATGATGCTTAGCCCTGG 1017

RESULT 11  
 AK020837  
 LOCUS  
 DEFINITION Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930009E21 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING  
 MARKS, full insert sequence.  
 AK020837  
 VERSION AK020837.1 GI:12861542  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carninci,P. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159  
 3  
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)  
 REFERENCE  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 REFERENCE  
 6  
 (bases 1 to 926)  
 AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hangaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
 COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for

further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGATCCCAAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using treatase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through two rounds of normalization to Rot = 20.0 and subtraction to Rot = 458.8. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTTAAATTAATCCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. Retina RNA was provided by Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.

Retina RNA was provided by Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.

#### FEATURES

Location/Qualifiers  
 1..926  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM.DB:A930009E21"  
 /db\_xref="MG1:1911678"  
 /db\_xref="taxon:10090"  
 /clone="A930009E21"  
 /tissue\_type="retina"  
 /clone\_libs="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stages="adult"  
 128..667  
 /note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING WARS (SPTR)AAL38196, evidence: FASTY, 100%ID, 100%length, match=777)  
 putative."  
 /codon\_start=1  
 /protein\_id="BAB32223.1"  
 /db\_xref="GI:12861543"  
 /translation="MPSVYAKVAHGWLYEGLSREKAEELLFPGFALIRESQT RRGCVSLVRLSRPASDRIRHYRIQLDNGWLYISPLTFPSLHALVEHYSELDGI CCLREPCVLQKLGPLPKDTPPVPTVPTSSLNWKLDRLSILFLEAPASGEASLSEGL LRESLSYISLAEDPLDDA"  
 907..912  
 /note="putative"  
 926  
 /note="putative"  
 polyA\_signal  
 polyA\_site

#### ORIGIN

Query Match 40.3%; Score 476.4; DB 11; Length 926;  
 Best Local Similarity 83.0%; Pred. No. 3.7e-109;  
 Matches 556; Conservative 0; Mismatches 111; Indels 3; Gaps 1;  
 QY 517 GGGCAGTTTCCCGGAGTGGCCGGCGAGCTGCTCGTACGACTCGGGGAGCCATTGAC 576  
 Db 4 GGGCAGTTTCCCGAGGTGAACAGGCAGACTATCTCTGAGACTCGGGGAGCCGCTGAC 63  
 QY 577 CATCGTCTCTCAGATGGAGACTGGTGGACGGTCTCTCTGAAGTCTCAGGCAGAGTA 636  
 Db 64 CATCATCTCTCAGGATGAGATTGGTGACAGTCCAGTCGGAAGTCTCAGGCAGAGTA 123  
 QY 637 TAAATCCCCAGCGTCCACGTCGGGCAAAAGTCTCCATGGGTGGTGTGTATGAGGCCCTGAG 696  
 Db 124 CCACATGCCAGTGTGTATGTGGCTAAAGTCGCCACCGGGTGGTGTGTACGAGGCCCTGAG 183  
 QY 697 CAGGAGAAACAGAGAACTGTGTGTACTCGGGAACCTCGAGGGGCGCTTCCTCAT 756  
 Db 184 CCGGAGAAACCGAGGAACCTACTCTGTACCTGGGAACCCCGGAGGGGCGCTTCCTCAT 243  
 QY 757 CCGGAGAGCCAGCAGGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCCCTGC 816



```
Query Match      38.2%; Score 452.2; DB 13; Length 660;
Best Local Similarity 83.4%; Pred. No. 3.9e-103;
Matches 538; Conservative 0; Mismatches 103; Indels 4; Gaps 2;

QY 517 GGGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGAC 576
DB 4 GGGCAGTTTCCCGCAGGTGACAGGCCAGCTATCTCTGAGACTCGGGAGCCGCTGAC 63

QY 577 CATGCTCTCTGAGATGGAGCTGGTGGACGGTCTGTGTAAGTCTCAGGCAGAGATGA 636
DB 64 CATCATCTCTGAGATGGAGATTTGGTGGACAGTCCAGTCCGAAGTCTCAGGCAGAGATGA 123

QY 637 TACATCCCCAGAGTCCACCTGGCGAAGCTCCCATGGGTGGCTGTATGAGGCCCTGAG 696
DB 124 CCACATGCCAGTGTGTATGTGGCTAAAGTCGCCACGGGTGGCTGTACAGGGCCCTGAG 183

QY 697 CAGGAGAAAGCAGAGGAATCTGTGTTTACCTGGGAACCTCTGGAGGGCCCTTCCTCAT 756
DB 184 CCGGGAGAAAGCCGAGGAATCTCTCTGTTTACCTGGGAACCCCGAGGGCCCTTCCTCAT 243

QY 757 CCGGGAGAGCCAGACAGGAGAGCTCTTACTCTCTGCTCAGTCCGCTCAGCCGCCCTGC 816
DB 244 CCGGGAGAGCCAGACAGGAGAGCTCTTACTCTCTGCTCAGTCCGCTCAGCCGCCCTGC 303

QY 817 ATCTCTGGGACCGGATCAGACACTACAGATCCACTGCTTGTGACAAATGGCTGGCTGTACAT 876
DB 304 ATCTTGGGACCGGATCAGACACTACAGATCCACTGCTTGTGACAAATGGCTGGCTGTACAT 363

QY 877 CTCACCGGCTCACCCTCCCTCACCCTCAGGAGCCCTGGTGACCATCTCTGAGCTGAC 936
DB 364 CTCACCTCGGCTCACCCTCCCTCACCCTCAGGAGCCCTGGTGACCATCTCTGAGCTGAC 423

QY 937 GGATGACATCTGTGCTCCTCAAGAGCCCTGTGCTCCTCAGAGGGCTGGCCCGCTCCC 996
DB 424 AGATGGCATCTGTGCTCCTCAGGAGCCCTGTGCTCCTCAGAGCTGGGCCCATACC 483

QY 997 TGGCAAGATATACCCCTACTGTGACTGTGAGAGGACACCACTCACTGGAAGAGCT 1056
DB 484 TGGCAAGATATACCCCTACTGTGACTGTGAGAGGACACCACTCACTGGAAGAGCT 543

QY 1057 GGACAGCTCCCTCTGTTTCTGAAG--CTGCCACAGGGGAGGAGTCTCTTCTCAGTGA 1113
DB 544 GGACCTGAGCTCTGTTTCTGAGAGCACTGCAAGTGGGAGGACATCTCTGCTATTGA 603

QY 1114 GGGTCTCCGGAGTCCCTCAGCTCTACATCAGCTGGAATGACGA 1158
DB 604 GGGGCTCCAGAT-CCTCACCTTCTACCTCATCCCGGTTGATGA 647

RESULT 13
LOCUS      BG284179          566 bp      mRNA      linear      EST 21-FEB-2001
DEFINITION 602408226F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520382 5',
mRNA sequence.
ACCESSION  BG284179
VERSION     BG284179.1 GI:13034866
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 566)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRAM10418 row: c column: 07
High quality sequence start: 2
High quality sequence stop: 566.
Location/Qualifiers
1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4520382"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/notes="Organ: prostate; vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match      34.2%; Score 405; DB 12; Length 566;
Best Local Similarity 100.0%; Pred. No. 3e-91;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 779 GGCTCTTACTCTGTCTGCTCAGTCCGCTCAGCCGCCCTGCATCTCGGACCGGATCAGACAC 838
DB 13 GGCTCTTACTCTGTCTGCTCAGTCCGCTCAGCCGCCCTGCATCTCGGACCGGATCAGACAC 72

QY 839 TACAGGATCCACTGCCTTGACAATGGCTGGCTGTACATCTCACCAGGCTCACCCTCCCC 898
DB 73 TACAGGATCCACTGCCTTGACAATGGCTGGCTGTACATCTCACCAGGCTCACCCTCCCC 132

QY 899 TCACCTCCAGGCGCTGTGGACCATTAATCTCTGAGCTGGCGGATGACATCTGCTGCTACTC 958
DB 133 TCACCTCCAGGCGCTGTGGACCATTAATCTCTGAGCTGGCGGATGACATCTGCTGCTACTC 192

QY 959 AAGAGCCCTGTGCTCTGCTCAGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCT 1018
DB 193 AAGAGCCCTGTGCTCTGCTCAGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCT 252

QY 1019 GTGACTGTGACAGAGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTTTCT 1078
DB 253 GTGACTGTGACAGAGACACCACTCACTGGAAGAGCTGGACAGCTCCCTCTGTTTCT 312

QY 1079 GAAGCTGCCACAGGGAGAGTCTTCTCTCAGTGGGGTCTCCGGAGTCCCTCAGCTTC 1138
DB 313 GAAGCTGCCACAGGGAGAGTCTTCTCTCAGTGGGGTCTCCGGAGTCCCTCAGCTTC 372

QY 1139 TACATCAGCCTGAATGACGAGGCTGTCTTCTTGGATGATGCTAG 1183
DB 373 TACATCAGCCTGAATGACGAGGCTGTCTTCTTGGATGATGCTAG 417

RESULT 14
LOCUS      AL844311          597 bp      mRNA      linear      EST 30-JUL-2002
DEFINITION AL844311 pool_yt_lib_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION  AL844311
VERSION     AL844311.1 GI:22019093
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 597)
Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
Sheridan,E.
Homo sapiens EST sequence
Unpublished (2002)
Contact: The Sanger Centre
The Sanger Centre
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
TITLE
JOURNAL
COMMENT
```

Email: humquery@sanger.ac.uk  
Sanger Centre name : sccdi0818.154136A  
Homo sapiens EST sequence. This sequence was generated as part of  
The Wellcome Trust Sanger Institute program to identify and  
annotate genes in the human genome. Incomplete or unconfirmed genes  
are experimentally analysed using a variety of cDNA library  
resources. This sequence was obtained from a PCR product generated  
from a pool of up to 100,000 cDNA clones derived from  
pool\_yt lib\_v SPD cDNA library. Further information can be found at  
<http://www.sanger.ac.uk/Teams/Team69/>.

## FEATURES

Location/Qualifiers  
1..597  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="20"  
/clone\_lib="pool\_yt lib\_v SPD"

## ORIGIN

Query Match 27.3%; Score 322.4; DB 9; Length 597;  
Best Local Similarity 99.7%; Pred. No. 2.3e-70;  
Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 860 AATGGCTGGCTGTACATCTCACCGGGCTCACTTCCCTCACTCCAGGCCCTGTGTGAC 919  
Db 13 AATGGCTGGCTGTACATCTCACCGGGCTCACTTCCCTCACTCCAGGCCCTGTGTGAC 72  
QY 920 CATTAATCTGAGCTGGGATGACATCTGTGCTCACTCAAGGAGCCCTGTGTGAC 979  
Db 73 CATTAATCTGAGCTGGGATGACATCTGTGCTCACTCAAGGAGCCCTGTGTGAC 132  
QY 980 AGGGCTGGCCCTCCCTGGCAAGATATACCCCTACCTGTGACGTGCGAGGACACCA 1039  
Db 133 AGGGCTGGCCCTCCCTGGCAAGATATACCCCTACCTGTGACGTGCGAGGACACCA 192  
QY 1040 CTCACCTGGAAGAGTGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAG 1099  
Db 193 CTCACCTGGAAGAGTGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAG 252  
QY 1100 TCTCTTCTCAGTGAGGCTCCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAG 1159  
Db 253 TCTCTTCTCAGTGAGGCTCCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAG 312  
QY 1160 GTGTCTCTTTGGATGATGCCTAG 1183  
Db 313 GTGTCTCTTTGGATGATGCCTAG 336

RESULT 15  
BU944126 960 bp mRNA linear EST 18-OCT-2002  
LOCUS AGENCOURT.10545003 NIH\_MGC.107 Homo sapiens cDNA clone  
DEFINITION IMAGE:6728350 5', mRNA sequence.  
ACCESSION BU944126  
VERSION BU944126.1 GI:24132945  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 960)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLCM3049 row: m column: 21  
High quality sequence stop: 628.

## FEATURES

Location/Qualifiers  
1..960  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:6728350"  
/issue\_type="adenocarcinoma, cell line"  
/lab\_host="DH103 (phage-resistant)"  
/clone\_lib="NIH\_MGC.107"  
/note="Osgan: breast; Vector: pOTB7; Site 1: EcoRI;  
Site 2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 27.1%; Score 321; DB 13; Length 960;  
Best Local Similarity 88.4%; Pred. No. 7.1e-70;  
Matches 381; Conservative 0; Mismatches 0; Indels 50; Gaps 1;  
QY 753 TCATCCGGAGAGCCAGACAGGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCC 812  
Db 1 TCATCCGGAGAGCCAGACAGGAGGCTCTTACTCTCTGTCACTCCGCTCAGCCGCC 60  
QY 813 CTGCATCTCGGAGCCGATCAGACATCAGATCCACTGCCCTTGACATGGCTGCTGT 872  
Db 61 CTGCATCTCGGAGCCGATCAGACATCAGATCCACTGCCCTTGACATGGCTGCTGT 120  
QY 873 ACATCTCACCGGCCCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTCT 932  
Db 121 ACATCTCACCGGCCCTCACCTTCCCTCACTCCAGGCCCTGTGTGACCATTTACTCT 176  
QY 933 TGGCGGATGACATCTGTGCTTACTCAAGAGCCCTGTGTCTGTGACAGGGCTGGCCGCC 992  
Db 177 -----GAGGGCTGGCCGCC 190  
QY 993 TCCTTGGCAAGGATATACCCCTACCTGTGACTGTGACAGGACACCACTCAACTGGAAAG 1052  
Db 191 TCCTTGGCAAGGATATACCCCTACCTGTGACTGTGACAGGACACCACTCAACTGGAAAG 250  
QY 1053 AGCTGGACAGCTCCCTCCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTG 1112  
Db 251 AGCTGGACAGCTCCCTCCCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTG 310  
QY 1113 AGGCTCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACAGGCTGTCTCTTTGG 1172  
Db 311 AGGCTCTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATGACAGGCTGTCTCTTTGG 370  
QY 1173 ATGATGCCTAG 1183  
Db 371 ATGATGCCTAG 381

Search completed: February 20, 2004, 01:41:31  
Job time : 7840.77 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:32:00 ; Search time 191.966 Seconds  
(without alignments)  
3419.919 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccaaggacc.....tctcttgatgatgctag 1183

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	10.4	2015	4	US-09-023-655-1105
2	119.4	10.1	2298	4	US-09-023-655-1158
3	110	9.3	2354	4	US-09-023-655-1080
4	107.4	9.1	2129	4	US-09-016-434-1452
5	101	8.5	675	1	US-08-707-793A-3
6	101	8.5	675	1	US-08-707-792A-3
7	92.6	7.8	2435	4	US-09-023-655-1313
8	91	7.7	2647	4	US-09-220-132-77
9	91	7.7	2647	5	PCT-US93-06251-77
10	90.8	7.7	1611	4	US-07-820-011A-3
11	90.8	7.7	1611	4	US-09-860-473-3
12	90.8	7.7	1611	5	PCT-US93-00445-3
13	85.6	7.6	1626	4	US-08-860-473-10
14	80.4	6.8	1602	1	US-07-820-011A-1
15	80.4	6.8	1602	5	PCT-US93-00445-1
16	80.4	6.8	1759	4	US-09-470-881-2
17	71	6.0	1491	2	US-09-006-675-1
18	71	6.0	1491	3	US-09-228-603A-1
19	68.4	5.8	282	2	US-09-006-675-5
20	68.4	5.8	282	3	US-09-228-603A-5
21	60.4	5.1	4517	4	US-09-470-881-7
22	60.4	5.1	4517	5	PCT-US93-06251-83
23	59.4	5.0	874	4	US-09-023-655-931
24	55.8	4.7	1467	4	US-09-579-182-2
25	55.8	4.7	1548	4	US-09-099-053-1
26	48	4.1	144	5	PCT-US93-06251-13
27	46	3.9	190	5	PCT-US93-06251-14

28	43.2	3.7	498	6	5219739-21	Patent No. 5219739
29	42.8	3.6	164	1	US-08-306-691B-28	Sequence 28, Appl
30	42.8	3.6	164	4	US-09-860-473-17	Sequence 17, Appl
31	42.8	3.6	164	5	PCT-US93-06251-70	Sequence 70, Appl
32	41.6	3.5	197	5	PCT-US95-10973A-18	Sequence 18, Appl
33	41.6	3.5	231	4	US-09-244-583-13	Sequence 13, Appl
34	41.6	3.5	444	4	US-09-392-932-6	Sequence 6, Appl
35	41.6	3.5	444	4	US-09-574-708A-1	Sequence 1, Appl
36	41.6	3.5	444	4	US-09-392-931-1	Sequence 1, Appl
37	41.6	3.5	456	5	PCT-US95-10973A-88	Sequence 88, Appl
38	41.6	3.5	467	5	PCT-US95-10973A-86	Sequence 86, Appl
39	41.6	3.5	473	3	US-08-718-904-1	Sequence 1, Appl
40	41.6	3.5	473	4	US-09-449-249-1	Sequence 1, Appl
41	41.6	3.5	473	5	PCT-US95-10973A-25	Sequence 25, Appl
42	41.6	3.5	495	4	US-09-244-583-25	Sequence 25, Appl
43	41.6	3.5	495	4	US-09-037-983C-14	Sequence 14, Appl
44	41.6	3.5	498	6	5194596-20	Patent No. 5194596
45	41.6	3.5	516	3	US-08-784-551C-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-023-655-1105  
; Sequence 1105, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1105:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2015 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g183911  
; US-09-023-655-1105

Query Match

10.4%; Score 123; DB 4; Length 2015;

[illegible]

RESULT 2  
US-09-023-655-1158  
; Sequence 1158, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Sellhammer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1158:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2298 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g187268  
US-09-023-655-1158

Query Match 10.1%; Score 119.4; DB 4; Length 2298;  
Best Local Similarity 53.3%; Pred. No. 1.8e-24;  
Matches 283; Conservative 0; Mismatches 236; Indels 12; Gaps 1;

QY 454 TGTCGAAGCCAGGGACCTGTGCACCATGGAAGCAGAGAGAAAGCAAGCCACACGCCCTGTGC 513  
DB 447 TTTACCTGGACAGAGTTCCTCAAACTAAAGATCCAGAGGAACAAGGAGACATTCGTGTAGC 506  
QY 514 CTTGGGCAGTTTCCCGCAGGTGGCCGCCGAGCTGTGCTGAGACTCGGGAGGCCATT 573  
DB 507 CTTGTACCCCTATGATGGCAATCCACCCGACGACTTGTCTTTCAAGAAAGGAGAGAAGAT 566  
QY 574 GACCATCGTCTGTGAGGATGGAGACTGTGGAGCGTGTCTGTGAAGTCTCAGGCGAGAGA 633  
DB 567 GAAAGTCCTGGAGCAGATGGAGAATGGTGGAAAGCAAAGTCCTTTTAAACAAAAAAGA 626  
QY 634 GTATACATCCCGACGTCACCTGGCGCAAGTCTCCC-----ATGGGTGGCT 681  
DB 627 AGCCTTACATCCCGACCACTATGTGCCCAACTCAACACTTAGAAACAGAGAGTGGTT 686  
QY 682 GTATAGGGCCTTAGCAGGAGAAAGCAGAGGAACCTGTGTTGTATTACCTGGGAACCCCTGG 741  
DB 687 TTTCAAGGATATAACCCAGGAAGACGACAGAAAGCGAGCTTTTGGCACCAAGAAATAGCGC 746  
QY 742 AGGGGCTTCCTCATCCGGGAGCCAGACAGGAGAGCTCTTACTCTCTGTCACTCCG 801  
DB 747 TGGAGCTTTCCTATTAGAGAAGTGAACATTTAAAGGAAGCTTCTCTCTGTCTGTCTCAG 806  
QY 802 CTTACGCGCCCTGTCATCTCGGACCGGATCAGACTACAGGATCCAATGCTTGCACAA 861  
DB 807 AGACTTTGACCCCTGTGCATGCTGATCTTATTAAAGCACTACAAAATTAGAAGTCTGATAA 866  
QY 862 TGCTCGCTGTACATCTCACCGGCTCACCTCCCTCACCTCCAGGCCCTGTGTGACCA 921  
DB 867 TGGGGCTATTACATCTCTCCAGAACTACTTTTCCCTGTATCAGCAATGATTAACA 926  
QY 922 TTACTCTGAGCTGGCGATGACATCTGTCCCTACTCAAGGAGCCCTGTGT 972  
DB 927 TTACCAAAAGCAGGCAGATGGCTTGTGCAAGAAATTGGAGAAGGCTTGTAT 977

RESULT 3  
US-09-023-655-1080  
Sequence 1080, Application US/09023655  
Patent No. 6607879  
GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.  
APPLICANT: Susan G. Stuart  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA

```

; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1080:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2354 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g182573
; US-09-023-655-1080

Query Match 9.3%; Score 110; DB 4; Length 2354;
Best Local Similarity 56.2%; Pred. No. 9.2e-22;
Matches 234; Conservative 0; Mismatches 170; Indels 12; Gaps 1;

QY 590 GATGAGACTGGTGCACGGTCTCTCAAGTCTCAGCAGAGATATACATCCCGAGC 649
DB 478 GAAGTGACTGGTGGAGGCTCGGTCTCTCAGCTCCGAAATCTGCGTGCATCCCGAGC 537

QY 650 GTCCACGTGGGCAAGT-----CTCCATGGTGGCTGTATGAGGGGCTGAGC 697
DB 538 AACTACGTGGCCCTGTTGACTCAATCCAGCTGAGAGTGTACTTGGAAAGATTGGG 597

QY 698 AGGGAAGACAGAGGAACCTGTTGTTACTCTGGAAACCTCGAGGGGCTTCTCATC 757
DB 598 AGAAGGATGCAGAGAGCAGCTGTTTCAACAGGCAACCCCGAGGGGGCTTCTCAT 657

QY 758 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGCA 817
DB 658 CGGGAAGCGAGACACCAAGAGTGCCTACTCTCTGTCCTCCGAGTGGATCAGACC 717

QY 818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGTCCTTGAATGCTGGCTGTACATC 877
DB 718 AGAGCGCATATGTGAGCATTTACAAGATCCGAAACTGGACATGGCGGCTACTACATC 777

QY 878 TCACCGCCCTCACTTCCCTCACTCAGCGCCCTGTTGAGCACTTACTCTGAGCTGGC 937
DB 778 ACCACAGGTTCACTTCACTCGTGCAGAGCTGTTGAGCACTACTGAGAGTGAAT 837

QY 938 GATGACATCTGCTGCTACTCAAGAGCCCTGTGTCTCTGAGAGGGTGGCCCGCT 993
DB 838 GACGGGCTGTGCAACCTGCTATCGCGCCCTGACCATCATGAAGCCGAGAGCT 893

RESULT 4
US-09-016-434-1452
; Sequence 1452, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer

```

```

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1452:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g775207
; US-09-016-434-1452

Query Match 9.1%; Score 107.4; DB 4; Length 2129;
Best Local Similarity 53.2%; Pred. No. 4.9e-21;
Matches 259; Conservative 0; Mismatches 216; Indels 12; Gaps 1;

QY 510 TGGCCCTGGGCAAGTTTCCCGCAGGTGGCCCGCGAGTGTGCTGAGACTCGGGGAGC 569
DB 259 TCGCTCTGCACAGCTATGAGCCCTCTCAGCAGGAGATCTGGGCTTGAGAAGGGGAAC 318

QY 570 CATTGACCATCTCTCTGAGGATGAGACTGTGTGACCGTGTCTGTGAAGTCTCAGGA 629
DB 319 AGCTCGCATCTCTGAGAGCAGCGCGAGTGTGTGAAAGCGCAGTCCCTGACCCGGGCC 378

QY 630 GAGAGTATAACATCCCGAGCGTCCACGTGGGCAAA-----GTCTCCCATGGGT 677
DB 379 AGAAGGCTTCATCCCTTCAATTTTGGCCAAAGCAGACAGCTTGGAGCCGACCT 438

QY 678 GCGTGTATCAGGGCTGAGCAGGAGAAAGCAGAGAACTGCTGTGTGTACTCTGGAACC 737
DB 439 GGTTCCTCAAGAACCTGAGCCGCAAGGACGCGAGCGCAGCTCTCTGCGCCCGGGAACA 498

QY 738 CTGGAAGGGGCGCTTCTCTCATCCGGGAGAGCCAGACAGGAGAGGCTTACTCTCTGTGAG 797
DB 499 CTCACGGCTCTCTCTCATCCGGGAGAGGAGAGCACCGCGGGATCGTTTTCACGTGTGG 558

QY 798 TCCGCTCAGCGCCCTCTGATCTCGGACCGGATCAGACACTACAGGATCCACTGCCCTTG 857
DB 559 TCCGGGACTTCGACCAGAACCCAGGAGAGGTGGTGAACATTATCAAGATCCGTAATCTGG 618

QY 858 ACAATGGTGGCTGTACATCTCAGCGCCCTCACTTCCCTCACTCCAGGCCCTGGTGG 917
DB 619 ACAAGGTGGCTTCTACATCTCCCTCGAATCACTTTTCCCGGCTGCTGATGATGTTCC 678

```

918 ACCATTACTCTGAGCTGGGGGATGACATCTCTGCTGCTACTCTAAGAGGCCCTGTGTCTCTGC 977  
 679 GGCATTACACCAATGCTTTTCAGATGGGCTGTGCACACGGTTGAGCGCGCCCTGCCAGACCC 738  
 978 AGAGGGC 984  
 739 AGAAGCC 745

```

1  RESULT 5
2  US-08-707-793A-3
3  ; Sequence 3, Application US/08707793A
4  ; Patent No. 5776696
5  ; GENERAL INFORMATION:
6  ; APPLICANT: SALOWE, SCOTT P.
7  ; TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING
8  ; TITLE OF INVENTION: FUSION PROTEINS
9  ; NUMBER OF SEQUENCES: 17
10 ; CORRESPONDENCE ADDRESS:
11 ; ADDRESSEE: Merck & Co., Inc.
12 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
13 ; CITY: Rahway
14 ; STATE: NJ
15 ; COUNTRY: USA
16 ; ZIP: 07065-0900
17 ; COMPUTER READABLE FORM:
18 ; MEDIUM TYPE: Diskette
19 ; COMPUTER: IBM Compatible
20 ; OPERATING SYSTEM: DOS
21 ; SOFTWARE: FastSeq for Windows Version 2.0
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/08/707,793A
24 ; FILING DATE: 04-SEP-1996
25 ; CLASSIFICATION: 435
26 ; PRIOR APPLICATION DATA:
27 ; APPLICATION NUMBER:
28 ; FILING DATE:
29 ; ATTORNEY/AGENT INFORMATION:
30 ; NAME: Camata, Valerie J
31 ; REGISTRATION NUMBER: 35,090
32 ; REFERENCE/DOCKET NUMBER: 19494
33 ; TELECOMMUNICATION INFORMATION:
34 ; TELEPHONE: 908-594-3902
35 ; TELEFAX: 908-594-4720
36 ; TELEX:
37 ; INFORMATION FOR SEQ ID NO: 3:
38 ; SEQUENCE CHARACTERISTICS:
39 ; LENGTH: 675 base pairs
40 ; TYPE: nucleic acid
41 ; STRANDEDNESS: single
42 ; TOPOLOGY: linear
43 ; MOLECULE TYPE: Genomic DNA
44 US-08-707-793A-3

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

QY 917 GACATTACTCTGAGTGGGGATGACATCTGCTGCTACTCAAGAGCCCTG 969  
|||||  
Db 613 CGCATTACCAATGCTTCAGATGGGCTGTGCACACGGTTGAGCGCCCTG 665

## RESULT 7

US-09-023-655-1313  
; Sequence 1313, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREMITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1313:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2435 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g338227  
; US-09-023-655-1313

Query Match 7.8%; Score 92.6; DB 4; Length 2435;  
Best Local Similarity 56.9%; Pred. No. 9.6e-17;  
Matches 170; Conservative 0; Mismatches 129; Indels 0; Gaps 0;  
QY 672 ATGGGTGGCTGATGAGGCTTGACGAGGAGAAAGCAGAGAACTGCTGTGTACCTG 731  
Db 810 AAGAGTGGTACTTGGAAACTTGGCCGAAAGATGCTGAGCAGCAGCTATTGCTCTTG 869  
QY 732 GGAACCTTGAGGGGCTTCTCTCATCCGGGAGAGCCAGCAGGAGAGGCTCTTACTCTC 791  
Db 870 GAAACCAAGAGGTACCTTCTTATCCGGAGAGTGAACCAACCAAGGGTCTATTAC 929  
QY 792 TGTAGTCCGCTCAGCGCCCTGATCTCTGGACCGGATCAGACACTACAGGATCCACT 851  
Db 930 TTCTATCCGTGATGGGATGATGAAGAGGAGACCATGTCAACACATTATAAATTGCA 989  
QY 852 GCCTTGACAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCC 911  
Db 990 AACTTGACAATGGTGAATACATATTACCACCGCGCCAGCTTTGAACACTTTCAGCAGC 1049

QY 912 TGGTGGACCATTAATCTGAGTGGCGGATGACATCTGCTGCTACTCAAGGAGCCCTGT 970  
|||||  
Db 1050 TTGTACAACATTAATCTGAGAGAGCTCAGGCTCTGCTGCCGCTAGTAGTTCCTGT 1108

## RESULT 8

US-09-220-132-77  
; Sequence 77, Application US/09220132  
; Patent No. 6506607  
; GENERAL INFORMATION:  
; APPLICANT: Shvjan, Andrew W.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 07334-074001  
; CURRENT APPLICATION NUMBER: US/09/220,132  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: US 60/079,303  
; PRIOR FILING DATE: 1997-12-24  
; PRIOR APPLICATION NUMBER: US 60/068,821  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 77  
; LENGTH: 2647  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-220-132-77

Query Match 7.7%; Score 91; DB 4; Length 2647;  
Best Local Similarity 56.5%; Pred. No. 2.9e-16;  
Matches 169; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 672 ATGGGTGGCTGATGAGGCTTGACGAGGAGGAGAGAAAGCAGAGAACTGCTGTCTTACCTG 731  
Db 1019 AAGAGTGGTACTTTGGAAACTTGGCGGAAAAGATGCTGAGCAGACTATTGCTCTTG 1078  
QY 732 GGAACCTTGAGGGGCTTCTCTCATCCGGGAGAGCCAGCAGGAGGCTCTTACTCTC 791  
Db 1079 GAAACCAAGAGGTACCTTCTTATCCGGGAGAGTGAACCAACCAAGGTGCTATTAC 1138  
QY 792 TGTGAGTCCGCTCAGCGCCCTGCATCTGGACCGGATCAGACACTACAGATCCACT 851  
Db 1139 TTCTATCCGTGATGGGATGATGAAGAGGAGACCATGTCAACACATTATAAATTGCA 1198  
QY 852 GCCTTGACAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCC 911  
Db 1199 AACTTGACAATGGTGGATACATACATACCAACCGGCGCCAGTTTGAACACTTCAGAGC 1258  
QY 912 TGGTGGACCATTAATCTGAGTGGCGGATGACATCTGCTGCTACTCAAGGAGCCCTGT 970  
Db 1259 TTGTACAACATTAATCTGAGAGAGCTGCAAGGTCTCTGCTGCCGCTAGTAGTTCCTGT 1317

## RESULT 9

PCT-US93-06251-77  
; Sequence 77, Application PC/TUS9306251  
; GENERAL INFORMATION:  
; APPLICANT: Wickstrom, Eric and Rife, Jason P.  
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digilio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2647 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-06251-77

Query Match          7.7%; Score 91; DB 5; Length 2647;
Best Local Similarity 56.5%; Pred. No. 2.9e-16;
Matches 169; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 672 ATGGTGGCTGTATGAGGGCTTCTCTATCGGAGCCAGGAGGCTCTTACTCTC 731
Db 1019 AAGAGTGTACTTTTGGAAAATTGGCCGAAAGATGCTGAGCGACAGCTATTGCTCTTG 1078
QY 732 GGAACCTTGGAGGGCTTCTCTATCGGAGCCAGGAGGCTCTTACTCTC 791
Db 1079 GAAACCAAGAGGTACTTTCTTATCGGAGTGAACCAACCAAGGTCCTATTAC 1138
QY 792 TGTGATCGGCTCAGCGGCTTCATCTTGGGACGGATCAGACACTACAGGATCCACT 851
Db 1139 TTTCTATCTCTGATTTGGATGATGATGAAGAGACCATGTCAACATTAATAAATTCGCA 1198
QY 852 GCCTTGACATGGTGGCTGTACTCTCAGCGGCTCCTCCCTCTACTCCAGGCC 911
Db 1199 AACTTGCAATGGTGGATCTACATTTACCACCGGCGCCAGTTTGAACACATTCAGCAGC 1258
QY 912 TGGTGGACCATTTACTCTGAGTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGT 970
Db 1259 TTGTACACATTTACTCAGAGAGCTCAGGCTCTCTGCTCGCGCTAGTAGTTCCCTGT 1317

RESULT 10
US-07-820-011A-3
; Sequence 3, Application US/07820011A
; Patent No. 5336615
; GENERAL INFORMATION:
; APPLICANT: Bell, Leonard
; APPLICANT: Madzi, Joseph A.
; APPLICANT: Warren, Stephen L.
; APPLICANT: Luthringer, Daniel J.
; TITLE OF INVENTION: Genetically Engineered
; TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced
; TITLE OF INVENTION: Migration
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb storage
; COMPUTER: IBM PC XT
; OPERATING SYSTEM: PC-DOS/MS-DOS 2.10
; SOFTWARE: Displaywrite 3
```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,011A
; FILING DATE: 19920106
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: LB-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255 1400
; TELEFAX: (203) 254 1101
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1611
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: No
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: Chromosome 20
; PUBLICATION INFORMATION:
; AUTHORS: Anderson, Stephen K.
; AUTHORS: Gibbs, Carol P.
; AUTHORS: Tanaka, Akio
; AUTHORS: Kung, Hsing-Jien
; AUTHORS: Fujita, Donald J.
; TITLE: Human Cellular src Gene:
; TITLE: Nucleotide Sequence and Derived Amino
; TITLE: Acid Sequence of the Region Coding for
; TITLE: the Carboxy-Terminal Two-Thirds of
; TITLE: pp60c-src
; JOURNAL: Molecular and Cellular Biology
; VOLUME: 5
; ISSUE: 5
; PAGES: 1122-1129
; DATE: May, 1985
; PUBLICATION INFORMATION:
; AUTHORS: Tanaka, Akio
; AUTHORS: Gibbs, Carol P.
; AUTHORS: Arthur, Richard R.
; AUTHORS: Anderson, Stephen K.
; AUTHORS: Kung, Hsing-Jien
; AUTHORS: Fujita, Donald J.
; TITLE: DNA Sequence Encoding the
; TITLE: Amino-Terminal Region of the Human c-src
; TITLE: Protein: Implications of Sequence
; TITLE: Divergence among src-Type Kinase
; TITLE: Oncogenes
; JOURNAL: Molecular and Cellular Biology
; VOLUME: 7
; ISSUE: 5
; PAGES: 1978-1983
; DATE: May, 1987
; US-07-820-011A-3
```

```

Query Match          7.7%; Score 90.8; DB 1; Length 1611;
Best Local Similarity 53.4%; Pred. No. 2.6e-16;
Matches 222; Conservative 0; Mismatches 182; Indels 12; Gaps 1;

QY 556 GAGACTCGGGAGCCCAATGACCATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGTC 615
Db 318 GAAAGGGGAGCGGCTCCAGATTGTCAACACACAGAGGAGACTGGTGGCGCCACTC 377
QY 616 TGAAGTCTCAGGAGAGATATACATCCCCAGCGTCCACGTGG-----GCAA 663
Db 378 GCTCAGCACAGGACAGACAGCGTACATCCCAGCACTACGTGGCGCCCTCCGACTCCAT 437
QY 664 AGTCTCCCATGGTGGCTGTATGAGGCGCTTGTAGGCGCTTGTAGGAGGAGAAAGCAGAGGAATGCTGTT 723
```

Db 438 CCAGGCTGAGGAGTGGTATTTTGGCAAGATCACAGACGGGAGTCAGACGGTACTGCT 497  
QY 724 GTTACTGGGAACCTCGGAGGCGCTTCTCATCCGGGAGAGCCAGACAGAGAGGCTC 783  
Db 498 CAATGACAGAACCCGAGAGGACCTTCTCTGCGAGAAAGTGAGACCAAGAGGTC 557  
QY 784 TTACTCTCTGTCAGTCCGCTCAGCGCCCTCAGCGCCCTCATCTGGGACCGGATCAGACACTACAG 843  
Db 558 CTACTGCCCTCTCAGTGTCTGACTTGTGACAAACGCGGAGGCTCAACGTGAAGCACTACAA 617  
QY 844 GATCCACTGCCCTTGACAAATGGTGGCTGTACATCTCACGGGCTCAGTCCCTCACT 903  
Db 618 GATCCGCAAGCTGACAGCGGCGCTTCTTACATCACTCCCGCACCCAGTTCACAGCCT 677  
QY 904 CCAGGCCCTGTGGACCAATATCTGTAGCTGGCGGATGACATCTGCTGCTACTCA 959  
Db 678 GCAGCAGCTGGTGGCGCTACTACTTCCAAACACGCGGATGGCTGTGCCACCGCTCA 733

## RESULT 11

US-09-860-473-3  
; Sequence 3, Application US/09860473  
; Patent No. 6656732  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Andrew T. Watt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-C EXPRESSION  
; FILE REFERENCE: RTS-0222  
; CURRENT APPLICATION NUMBER: US/09/860,473  
; CURRENT FILING DATE: 2001-05-18  
; NUMBER OF SEQ ID NOS: 169  
; SEQ ID NO 3  
; LENGTH: 1611  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1611)  
US-09-860-473-3

Query Match 7.7%; Score 90.8; DB 4; Length 1611;  
Best Local Similarity 53.4%; Pred. No. 2.6e-16;  
Matches 222; Conservative 0; Mismatches 182; Indels 12; Gaps 1;

QY 556 CAGATCTCGGGAGGCAATTGACCATCTCTGTGAGGTGGAGACTGTGGAGGTCGTGTC 615  
Db 318 GAAAGGCGAGCGGCTCCAGATTGTCAACACAGAGAGGAGACTGGTGGCCCACTC 377  
QY 616 TGAAGTCTCAGCAGAGAGATATACATCCCGAGCGTCCACGTGG-----GCAA 663  
Db 378 GCTCAGCAGACAGCAGACAGGCTACATCCCGAGCAACTAGTGGCGCCCTCCGACTCCAT 437  
QY 664 AGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGCAGGAGAGACGAGAACTGCTGTT 723  
Db 438 CCAGGCTCAGAGTGTATTTTGGCAAGATCACAGACGGGAGTCAGAGCGGTACTGCT 497  
QY 724 GTTACTCTGGGAACCTTGAGGGGCTTCTCATCCGGGAGAGCCAGACAGGAGGCTC 783  
Db 498 CAATGACAGAAACCCGAGAGGACCTTCTCTGTCGAGAAAGTCAGACACAGAAAGGTC 557  
QY 784 TTACTCTGTGCTCGGCTCAGCGCCCTGATCTCTGGGACCGGATGACAGACTACAG 843  
Db 558 CTACTGCTCTCAGTGTCTGACTTGTGACAAACGCGGAGGCTCAACGTGAAGCACTACAA 617  
QY 844 GATCCACTGCTTGACAAATGGCTGGCTGTATCTCATCCCGCGCTCAGCTTCCCTCACT 903  
Db 618 GATCCGCAAGCTGACAGCGGCGCTTCTACATCACTCCCGCACCCAGTTCACAGCCT 677  
QY 904 CCAGGCCCTGTGACCAATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTCA 959  
Db 678 GCAGCAGCTGGTGGCGCTACTACTTCCAAACACGCGGATGGCTGTGCCACCGCTCA 733

## RESULT 12

PCT-US93-00445-3  
; Sequence 3, Application PC/TUS9300445  
; GENERAL INFORMATION:  
; APPLICANT: Bell, Leonard  
; APPLICANT: Madri, Joseph A.  
; APPLICANT: Warren, Stephen L.  
; APPLICANT: Luthringer, Daniel J.  
; TITLE OF INVENTION: Genetically Engineered  
; TITLE OF INVENTION: Endothelial Cells  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maurice M. Klee  
; STREET: 1951 Burr Street  
; CITY: Fairfield  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 06430  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 760 Kb storage  
; COMPUTER: DELL 486/50  
; OPERATING SYSTEM: DOS 5.0  
; SOFTWARE: Displaywrite 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/00445  
; FILING DATE: 19930105  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/820,011  
; FILING DATE: 06-JAN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Klee, Maurice M.  
; REGISTRATION NUMBER: 30,399  
; REFERENCE/DOCKET NUMBER: ALX-101PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (203) 255 1400  
; TELEFAX: (203) 254 1101  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1611  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: Double  
; TOPOLOGY: Linear  
; MOLECULE TYPE: CDNA to mRNA  
; HYPOTHETICAL: No  
; ANTI-SENSE: No  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapien  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT: Chromosome 20  
; PUBLICATION INFORMATION:  
; AUTHORS: Anderson, Stephen K.  
; AUTHORS: Gibbs, Carol P.  
; AUTHORS: Tanaka, Akio  
; AUTHORS: Kung, Hsing-Jien  
; AUTHORS: Fujita, Donald J.  
; TITLE: Human Cellular src Gene:  
; TITLE: Nucleotide Sequence and Derived Amino  
; TITLE: Acid Sequence of the Region Coding for  
; TITLE: the Carboxy-Terminal Two-Thirds of  
; TITLE: pp60c-src  
; JOURNAL: Molecular and Cellular Biology  
; VOLUME: 5  
; ISSUE: 5  
; PAGES: 1122-1129  
; DATE: May, 1985  
; PUBLICATION INFORMATION:  
; AUTHORS: Tanaka, Akio  
; AUTHORS: Gibbs, Carol P.  
; AUTHORS: Arthur, Richard R.  
; AUTHORS: Anderson, Stephen K.  
; AUTHORS: Kung, Hsing-Jien  
; AUTHORS: Fujita, Donald J.

; TITLE: DNA Sequence Encoding the  
 ; TITLE: Amino-Terminal Region of the Human c-src  
 ; TITLE: Protein: Implications of Sequence  
 ; TITLE: Divergence among src-Type Kinase  
 ; TITLE: Oncogenes  
 ; JOURNAL: Molecular and Cellular Biology  
 ; VOLUME: 7  
 ; ISSUE: 5  
 ; PAGES: 1978-1983  
 ; DATE: May, 1987  
 ; PCT-US93-00445-3

Query Match 7.7%; Score 90.8; DB 5; Length 1611;  
 Best Local Similarity 53.4%; Pred. No. 2.6e-16;  
 Matches 222; Conservative 0; Mismatches 182; Indels 12; Gaps 1;

QY	556	GAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTGTC	61.5
DB	318	GAAGGGAGCGGCTCCAGATTGTCAACAACAGAGGAGACTGGTGGTGGCCCACTC	377
QY	616	TGAAGTCTCAGCAGAGAGTATACATCCCGAGCGTCCAGTGG-----GCAA	663
DB	378	GCTCAGCAGCAGCAGCAGAGCGTATATCCCGCAGCACTAGCTGGCGCCCTCCGACTCCAT	437
QY	664	AGTCTCCCATGGGTGGTGTATGAGGGCCTCAGCAGGAGAGAAACAGAGAACTGCTGTT	723
DB	438	CCAGGCTGAGGAGTGGTATTTTGGCAAGATCACAGAGGAGTCAAGCGGTACTGCT	497
QY	724	GTTACCTGGGAACCTGGAGGGGCTTCTTATCCGGAGAGCAGACCGAGAGAGGCTC	783
DB	498	CAATGCAGAGAACCCGAGAGGAGGACCTTCTCTGTCGAGAAAGTGAGACCACGAAAGGTGC	557
QY	784	TTACTCTCTGTCAGTCCGCGCTCAGCCGCTGTCATCTCGGACCGGATCAGACACTACAG	843
DB	558	CTACTGCTCTCAGTGTCTGACTTTCGACAGCCGAGGGCTCAACGTGAAGCACTACAA	617
QY	844	GATCCACTGCTTGACAAATGCTGGCTGTATATCTACCGCGCTCACCCTCCCTCACT	903
DB	618	GATCCGAAAGTGGACAGCGCGGCTTCTATCATCCTCCCGCACCAGTTCACAGCCT	677
QY	904	CCAGGCGCTGTGACCACTACTCTGAGTGGGGATGACATCTGTCGCTACTCA 959	
DB	678	GCACAGCTGTGCTTACTTACTTCCAAACAGCGCGATGGCTGTGCCACCGGCTCA 733	

RESULT 13  
 US-09-860-473-10  
 ; Sequence 10, Application US/09860473  
 ; Patent No. 6656732  
 ; GENERAL INFORMATION:  
 ; APPLICANT: C. Frank Bennett  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-C EXPRESSION  
 ; FILE REFERENCE: RTS-0222  
 ; CURRENT APPLICATION NUMBER: US/09/860,473  
 ; NUMBER OF SEQ ID NOS: 169  
 ; SEQ ID NO 10  
 ; LENGTH: 1626  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(1626)  
 US-09-860-473-10

Query Match 7.6%; Score 89.6; DB 4; Length 1626;  
 Best Local Similarity 56.9%; Pred. No. 5.8e-16;  
 Matches 164; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY	672	ATGGGTGGCTGTATGAGGGCTGAGCAGGAGAGAAACAGAGAGTGGTGTGTACTG 731
DB	461	AGGAGTGGTACTTTGGCAAGATCACTAGACGGGAATCAGAGCGGCTGCTCAACGGCG 520

QY	732	GGAACTCTGAGGGGCTTCTCTATCCGGGAGAGCCAGACAGAGAGCTCTTACTCTC 791
DB	521	AGAACTCTGAGGGGCTTCTCTGAGGGAGAGTGAACCAACAAAGGTGCTTACTGCC 580
QY	792	TGTCTAGTCCGCTCAGCCGCTCTGATCTCTGGGACCGGATCAGACACTACAGATCCACT 851
DB	581	TCTCTGATCCGACTTCGACATCCAGGGGCTAAATGTGAACTACAGATCCGCA 640
QY	852	GCTTGTGACATGGCTGGCTGTATCTACCTCAGCGGCTCACCCTCCCTCCTCCTCAGGGCC 911
DB	641	AGCTGACAGCGCGGTTTCTATCATCCTCCGACCCAGTTCACAGCTTCGACGAGC 700
QY	912	TGCTGACCACTTACTCTGAGCTGGGGATGACATCTGCTGCTACTCA 959
DB	701	TGCTGCTTACTTCTTCCAAACATGCTGATGGCTGTGTCACCGGCTCA 748

RESULT 14  
 US-07-820-011A-1  
 ; Sequence 1, Application US/07820011A  
 ; Patent No. 5336615  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell, Leonard  
 ; APPLICANT: Madri, Joseph A.  
 ; APPLICANT: Warren, Stephen L.  
 ; APPLICANT: Luthringer, Daniel J.  
 ; TITLE OF INVENTION: Genetically Engineered  
 ; TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced  
 ; TITLE OF INVENTION: Migration  
 ; TITLE OF INVENTION: and Plasminogen Activator Activity  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Maurice M. Klee  
 ; STREET: 1951 Burr Street  
 ; CITY: Fairfield  
 ; STATE: Connecticut  
 ; COUNTRY: USA  
 ; ZIP: 06430  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 5.25 inch, 360 Kb storage  
 ; COMPUTER: IBM PC XT  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS 2.10  
 ; SOFTWARE: Displaywrite 3  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/820,011A  
 ; FILING DATE: 19920106  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Klee, Maurice M.  
 ; REGISTRATION NUMBER: 30,399  
 ; REFERENCE/DOCKET NUMBER: LB-101  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (203) 255 1400  
 ; TELEFAX: (203) 254 1101  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1602 base pairs  
 ; TYPE: NUCLEIC ACID  
 ; STRANDEDNESS: Double  
 ; TOPOLOGY: Linear  
 ; MOLECULE TYPE: cDNA to mRNA  
 ; HYPOTHETICAL: No  
 ; ANTI-SENSE: No  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Gallus, gallus  
 ; PUBLICATION INFORMATION:  
 ; AUTHORS: Takeya, Tatsuo  
 ; AUTHORS: Hanafusa, Hidesaburo  
 ; TITLE: Structure and Sequence of the  
 ; TITLE: Cellular Gene Homologous to the RSV src  
 ; TITLE: Gene and the Mechanism for Generating the  
 ; TITLE: Transforming Virus



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:23:34 ; Search time 2843.55 Seconds  
(without alignments)  
1456.787 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctcaaggacc.....tctcttgatgagctag 1183

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	784.4	66.3	786	14	US-10-043-649-1
2	775.4	65.5	864	10	US-09-814-353-21302
3	758.2	64.1	763	9	US-09-867-550-953
4	348	28.4	444	9	US-09-867-550-951
5	341	28.8	875	9	US-09-867-550-1915
6	213.4	18.0	320	10	US-09-814-353-17314
7	157.4	13.3	2665	9	US-09-954-456-499
8	157.4	13.3	3756	13	US-10-002-600-91
9	141.8	12.0	432	9	US-09-864-761-2829
10	141.8	12.0	448	9	US-09-864-761-15513
11	131.6	11.1	152	10	US-09-814-353-4631
12	131.6	11.1	152	10	US-09-814-353-10930
13	124.6	10.5	2343	15	US-10-062-674-2038
14	123	10.4	1924	15	US-10-193-720-1
15	123	10.4	2015	9	US-09-954-456-1983

16	123	10.4	2015	14	US-10-007-010-3	Sequence 3, Appl
17	123	10.4	2341	14	US-10-252-157-140	Sequence 140, App
18	119.4	10.1	2298	14	US-10-175-523-50	Sequence 50, Appl
19	119.4	10.1	2298	15	US-10-159-563-343	Sequence 343, App
20	110.6	9.3	2032	15	US-10-366-288-27	Sequence 27, Appl
21	110	9.3	1590	15	US-10-085-117-18	Sequence 18, Appl
22	110	9.3	2354	9	US-09-967-768A-300	Sequence 300, App
23	110	9.3	2354	15	US-10-353-690-123	Sequence 123, App
24	110	9.3	2354	15	US-10-085-117-17	Sequence 17, Appl
25	110	9.3	2343	14	US-10-240-965-114	Sequence 114, App
26	107.4	9.1	2017	15	US-10-062-674-1776	Sequence 1776, App
27	107.4	9.1	2129	10	US-09-960-706-954	Sequence 954, App
28	107.4	9.1	2129	15	US-10-305-720-1452	Sequence 1452, App
29	102.6	8.7	1911	9	US-09-917-800A-1611	Sequence 1611, App
30	99	8.4	1554	15	US-10-085-117-15	Sequence 15, Appl
31	99	8.4	2179	15	US-10-085-117-14	Sequence 14, Appl
32	96	8.1	96	9	US-09-864-761-19612	Sequence 19612, A
33	91	7.7	549	10	US-09-918-995-25044	Sequence 25044, A
34	91	7.7	1609	9	US-09-771-161A-30	Sequence 30, Appl
35	91	7.7	1995	9	US-09-771-161A-31	Sequence 31, Appl
36	91	7.7	4414	14	US-10-101-510-512	Sequence 512, App
37	89.6	7.6	2451	9	US-09-771-161A-4	Sequence 4, Appl
38	84.8	7.2	488	10	US-09-918-995-19745	Sequence 19745, A
39	83.8	7.1	2173	15	US-10-094-749-1087	Sequence 1087, App
40	76	6.4	342	15	US-10-062-674-1242	Sequence 1242, App
41	73.2	6.2	486	9	US-09-796-692-7419	Sequence 7419, App
42	73.2	6.2	486	14	US-10-040-862-7419	Sequence 7419, App
43	73.2	6.2	486	15	US-10-057-475B-7419	Sequence 7419, App
44	73.2	6.2	486	15	US-10-154-884B-7419	Sequence 7419, App
45	60.4	5.1	4343	15	US-10-153-563-184	Sequence 184, App

## ALIGNMENTS

### RESULT 1

US-10-043-649-1  
; Sequence 1, Application US/10043649  
; Publication No. US20030059924A1  
; GENERAL INFORMATION:  
; APPLICANT: Holland, Sacha J.  
; APPLICANT: Mendenhall, Marcy K.  
; APPLICANT: Pardo, Jorge  
; APPLICANT: Spencer Collin  
; APPLICANT: Fu, C. Alan  
; APPLICANT: Luo, Ying  
; APPLICANT: Payan, Donald G.  
; APPLICANT: Mancebo, Helena S.Y.  
; APPLICANT: Wu, Jun  
; APPLICANT: Zhou, Xiulan  
; APPLICANT: Shen, Mary  
; APPLICANT: Liao, X. Charlene  
; APPLICANT: Sheng, Ning  
; TITLE OF INVENTION: Cloning of a No. US20030059924A1  
; TITLE OF INVENTION: Retroviral-based Functional Screen  
; FILE REFERENCE: A-70219-1/RMS/DHR  
; CURRENT APPLICATION NUMBER: US/10/043,649  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/260,953  
; PRIOR FILING DATE: 2001-01-10  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 786  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(786)  
; OTHER INFORMATION:  
US-10-043-649-1

Query Match 66.3%; Score 784.4; DB 14; Length 786;

```

Best Local Similarity 99.9%; Pred. No. 1.5e-229;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGGAAGTGTGCGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTCTGTC 457
Db 1 ATGGGAAGTGTGCGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTCTGTC 60
QY 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 517
Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 518 GCGAGTTTCCCGGCGAGGTGGCCCGGCGAGGTGTCGTGAGACTCGGGGAGCCATTGACC 577
Db 121 GCGAGTTTCCCGGCGAGGTGGCCCGGCGAGGTGTCGTGAGACTCGGGGAGCCATTGACC 180
QY 578 ATCGTCTCTGAGGATGAGAGCTGGTGGACGTGCTGTCGAGTCTGAGAGAGAGATAT 637
Db 181 ATCGTCTCTGAGGATGAGAGCTGGTGGACGTGCTGTCGAGTCTGAGAGAGAGATAT 240
QY 638 AACATCCCAGCGTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTCGAGC 697
Db 241 AACATCCCAGCGTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTCGAGC 300
QY 698 AGGGAAGAGCAGAGAGACTCTGTGTTACTCTGGAACTCTGAGAGGGGCTTCTTCATC 757
Db 301 AGGGAAGAGCAGAGAGACTCTGTGTTACTCTGGAACTCTGAGAGGGGCTTCTTCATC 360
QY 758 CGGAGAGCCAGACAGCAGAGAGGCTTACCTCTGTCAGTCCGCTCAGCCGCCCTGCA 817
Db 361 CGGAGAGCCAGACAGCAGAGAGGCTTACCTCTGTCAGTCCGCTCAGCCGCCCTGCA 420
QY 818 TCCTGGGACCGATCAGACACTCAGATCCACTGCGCTTGACAAATGGCTGGCTGTATC 877
Db 421 TCCTGGGACCGATCAGACACTCAGATCCACTGCGCTTGACAAATGGCTGGCTGTATC 480
QY 878 TCACCGGCGCTCACCTTCCCTCACTCCAGGCGCTGTGGACCACTTACTCTGAGCTGGG 937
Db 481 TCACCGGCGCTCACCTTCCCTCACTCCAGGCGCTGTGGACCACTTACTCTGAGCTGGG 540
QY 938 GATGACATCTGCTGCTACTCAAGAGCGCTGTGCTCTGAGAGGGGTGCCCGCTCCCT 997
Db 541 GATGACATCTGCTGCTACTCAAGAGCGCTGTGCTCTGAGAGGGGTGCCCGCTCCCT 600
QY 998 GCGAAGATATACCCCTACCTGTGACGTGTGAGAGGACACACTCACTGGAAGAGCTG 1057
Db 601 GCGAAGATATACCCCTACCTGTGACGTGTGAGAGGACACACTCACTGGAAGAGCTG 660
QY 1058 GACAGTCCCTCTGCTTTTGAAGCTGCCAAGGGAGGAGTCTCTCTCAGTGAGGGT 1117
Db 661 GACAGTCCCTCTGCTTTTGAAGCTGCCAAGGGAGGAGTCTCTCTCAGTGAGGGT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACAGCTGATGACGAGGCTGCTCTTTGGATGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACAGCTGATGACGAGGCTGCTCTTTGGATGAT 780
QY 1178 GCCTAG 1183
Db 781 GCCTAG 786

```

```

RESULT 2
US-09-814-353-21302
; Sequence 21302, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353

```

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21302
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 32, 862, 863, 864
; OTHER INFORMATION: n = A,T,C or G
; US-09-814-353-21302

```

```

Query Match 65.5%; Score 775.4; DB 10; Length 864;
Best Local Similarity 99.2%; Pred. No. 8.7e-227;
Matches 779; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GTGAGCTTCAAGACCCCAAGCCCTGCTGTCTGTGACAGAGCTCAAAGGGCCCTTGGGC 61
Db 54 GTGAGCTTCAAGACCCCAAGCCCTGCTGTCTGTGACAGAGCTCAAAGGGCCCTTGGGC 113
QY 62 CTTCCCTCCTCGCTCGCTGTGCTTGGAGGCTTCCCGAGTCCAGATCCCTTAGGAGC 121
Db 114 CTTCCCTCCTCGCTCGCTGTGCTTGGAGGCTTCCCGAGTCCAGATCCCTTAGGAGC 173
QY 122 ATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCTA 181
Db 174 ATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCTA 233
QY 182 CCGAAGCAGACACCTGAGCTCTCCCTGAGATCTCCCGAGCTCCAGAGTCTTGGGTGT 241
Db 234 CCGAAGCAGACACCTGAGCTCTCCCTGAGATCTCCCGAGCTCCAGAGTCTTGGGTGT 293
QY 242 CTTAGGACCAAGGACACTGGCAGACTTCCAGAAGGGGCCCAAGGCCCTAACCTGTCCAG 301
Db 294 CTTAGGACCAAGGACACTGGCAGACTTCCAGAAGGGGCCCAAGGCCCTAACCTGTCCAG 353
QY 302 CCAGAGCATGGCTCTCAGCAGAGCTGTCTTCCAGAGCTTTGATGACAAATTTCCC 361
Db 354 CCAGAGCATGGCTCTCAGCAGAGCTGTCTTCCAGAGCTTTGATGACAAATTTCCC 413
QY 362 TCGATGATGTGCTTCTGAGTGTCTCTGAGGAACTTGGGAAGTCTGCCCGAGAGAGA 421
Db 414 TCGATGATGTGCTTCTGAGTGTCTCTGAGGAACTTGGGAAGTCTGCCCGAGAGAGA 473
QY 422 AAATCTCTGCAAGCCCAAGCTTGAATCTCTCTGTCGAGGAGGAGCTGTGACCATG 481
Db 474 AAATCTCTGCAAGCCCAAGCTTGAATCTCTCTGTCGAGGAGGAGCTGTGACCATG 533
QY 482 GAAGCAGAGAGAAGCAAGGCCACAGCGCTGGGCGCTTCCCGGAGGTGGGCCG 541
Db 534 GAAGCAGAGAGAAGCAAGGCCACAGCGCTGGGCGCTTCCCGGAGGTGGGCCG 593
QY 542 CCGGAGCTGTGCTGAGACTCGGGAGCCATGACCATGCTCTCTGAGGATGAGAGATGG 601
Db 594 CCGGAGCTGTGCTGAGACTCGGGAGCCATGACCATGCTCTCTGAGGATGAGAGATGG 653
QY 602 TGGAGGCTGTGCTGAGAGTCTCAGGAGAGATATACATCCCGAGCTCCAGCTGGC 661
Db 654 TGGAGGCTGTGCTGAGAGTCTCAGGAGAGATATACATCCCGAGCTCCAGCTGGC 713

```

QY 662 AAGTCTCCATGGGTGGTGGTGTATGAGGGCTGACGAGGGAAGAGGAGAACTGCTG 721  
 Db 714 AAGTCTCCATGGGTGGTGGTGTATGAGGGCTGACGAGGGAAGAGGAGAACTGCTG 773  
 QY 722 TTGTTACCTGGGAACCTCTGGAGGGGCTTCTCATCCGGAGAGCCAGCAGAGGC 781  
 Db 774 TTGTTACCTGGGAACCTCTGGAGGGGCTTCTCATCCGGAGAGCCAGCAGAGGC 833  
 QY 782 TCTTA 786  
 Db 834 TCTTA 838

## RESULT 3

US-09-867-550-953  
 ; Sequence 953, Application US/09867550  
 ; Patent No. US20020082206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad,  
 ; APPLICANT: Conley, Pamela  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
 ; TITLE OF INVENTION: Thereby  
 ; FILE REFERENCE: 21402-013 (Cura-313)  
 ; CURRENT APPLICATION NUMBER: US/09/867,550  
 ; CURRENT FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: US2001-09-20  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 953  
 ; LENGTH: 763  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-867-550-953

Query Match 64.1%; Score 758.2; DB 9; Length 763;  
 Best Local Similarity 99.6%; Pred. No. 1.5e-221;  
 Matches 760; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 113 CTAAGAGATGGGGAGTGTATGATCCCTGGTGTGTAACAATGCTGCTGACGACAGAT 172  
 Db 1 CTAAGAGATGGGGAGTGTATGATCCCTGGTGTGTAACAATGCTGCTGACGACAGAT 60  
 QY 173 GCTGAGCTACCCAAACCAACACCTAGCTCTCCCTGAAGATCCTCCAGGCTGAGAGT 232  
 Db 61 GCTGAGCTACCCAAACCAACACCTAGCTCTCCCTGAAGATCCTCCAGGCTGAGAGT 120  
 QY 233 TGTGGTGTCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTAA 292  
 Db 121 TGTGGTGTCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCAAAGCCCTAA 180  
 QY 293 CTTGTCAGCAGAGCATGCTGCTCAGCAGAGCTGCTTCCCAAGCTTTGATGACAAAC 352  
 Db 181 CTTGTCAGCAGAGCATGCTGCTCAGCAGAGCTGCTTCCCAAGCTTTGATGACAAAC 240  
 QY 353 CAATTTCCCTCGATGATGCTTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCC 412  
 Db 241 CAATTTCCCTCGATGATGCTTCTGAGTGTCTGCTGAGGAAACAATGGGAAGTCTGCC 300  
 QY 413 AGCAGAAGAAATCTCTGCAAGCCCAAGCTTGTAGTCTCTGTCGAAGGCCAGGACCT 472  
 Db 301 AGCAGAAGAAATCTCTGCAAGCCCAAGCTTGTAGTCTCTGTCGAAGGCCAGGACCT 360  
 QY 473 GTACCATGGAAGCAGAGAGAGAGGACCAAGCCGTGGCCCTGGGAGTTCCTCCGGCA 532  
 Db 361 GTACCATGGAAGCAGAGAGAGAGGACCAAGCCGTGGCCCTGGGAGTTCCTCCGGCA 420  
 QY 533 GGTGGCCCGGCGAGTGTCTGAGACTCGGGAGCCATTTGACCATGCTCTCTGAGGAT 592  
 Db 421 GGTGGCCCGGCGAGTGTCTGAGACTCGGGAGCCATTTGACCATGCTCTCTGAGGAT 480

QY 593 GGAGACTGGTGGAGGGTGTCTCTGAAATCTCAGGACGAGAGTATACATCCCAGGCTC 652  
 Db 481 GGAGACTGGTGGAGGGTGTCTCTGAAATCTCAGGACGAGAGTATACATCCCAGGCTC 540  
 QY 653 CACGTGGGCAAGTCTCCATGGGTGGTGTATGAGGGCTGAGCAGGGAGAGAGAGAG 712  
 Db 541 CACGTGGGCAAGTCTCCATGGGTGGTGTATGAGGGCTGAGCAGGGAGAGAGAGAG 600  
 QY 713 GAATCTGTGTTTACCTGGGAACCTCTGGAGGGGCTTCTCTCATCCGGAGAGCCAGACC 772  
 Db 601 GAATCTGTGTTTACCTGGGAACCTCTGGAGGGGCTTCTCTCATCCGGAGAGCCAGACC 660  
 QY 773 AGGAGAGGCTCTTACTCTCTCTGAGTCCGCTCAGCCGCTGATCCTCTGGACCCGATC 832  
 Db 661 AGGAGAGGCTCTTACTCTCTCTGAGTCCGCTCAGCCGCTGATCCTCTGGACCCGATC 720  
 QY 833 AGACACTACAGGATCCACTGCTTGAACAATGGTGGTGTGTACA 875  
 Db 721 AGACACTACAGGATCCACTGCTTGAACAATGGTGGTGTGTACA 763

## RESULT 4

US-09-867-550-951  
 ; Sequence 951, Application US/09867550  
 ; Patent No. US20020082206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad,  
 ; APPLICANT: Conley, Pamela  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
 ; TITLE OF INVENTION: Thereby  
 ; FILE REFERENCE: 21402-013 (Cura-313)  
 ; CURRENT APPLICATION NUMBER: US/09/867,550  
 ; CURRENT FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: US2001-09-20  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 951  
 ; LENGTH: 444  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-867-550-951

Query Match 29.4%; Score 348; DB 9; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 4e-96;  
 Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 CTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTTAACCTGTCCAG 301  
 Db 1 CTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTTAACCTGTCCAG 60  
 QY 302 CCAGAGCATGCTCTCAGCAGAGCTGCTTCCCAAGCTTTGATGACAAACCAATTTCCC 361  
 Db 61 CCAGAGCATGCTCTCAGCAGAGCTGCTTCCCAAGCTTTGATGACAAACCAATTTCCC 120  
 QY 362 TCGATGATGCTCTCTGAGTGTCTCTGAGGAAACAATGGGAAGTCTGCCAGAGAAGA 421  
 Db 121 TCGATGATGCTCTCTGAGTGTCTCTGAGGAAACAATGGGAAGTCTGCCAGAGAAGA 180  
 QY 422 AAATCTCTGCAAGCCCAAGCTTGTAGTTCCTCTGTCGAAGGCCAGGACCTGTGACCATG 481  
 Db 181 AAATCTCTGCAAGCCCAAGCTTGTAGTTCCTCTGTCGAAGGCCAGGACCTGTGACCATG 240  
 QY 482 GAAGCAGAGAGAGCAAGCCGACAGCCGTGGCCCTGGGAGTTCCTCCGGAGGTGGCCCG 541  
 Db 241 GAAGCAGAGAGAGCAAGCCGACAGCCGTGGCCCTGGGAGTTCCTCCGGAGGTGGCCCG 300  
 QY 542 GCCGAGTGTCTGCTGAGACTCGGGAGCCATTGACCATCTCTCTGAG 589

Db 301 GCCGAGCTGCTCGTCGAGACTCGGGGAGCCATTGACCATCGTCTCTGAG 348

## RESULT 5

US-09-867-550-1915  
; Sequence 1915, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehrahan, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; FILE REFERENCE: 21402-013 (Cura-3113)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; PRIOR FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1915  
; LENGTH: 875  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)  
; OTHER INFORMATION: Wherein n is one of a or t or c or g  
US-09-867-550-1915

Query Match 28.8%; Score 341; DB 9; Length 875;  
Best Local Similarity 100.0%; Pred. No. 6.5e-94;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	843	GGATCCACTGCTGACAAATGGCTGGGTGATACATCTCACCGCGCTCACCTTCCCTCTAC	902
Db	2	GGATCCACTGCTGACAAATGGCTGGGTGATACATCTCACCGCGCTCACCTTCCCTCTAC	61
QY	903	TCCAGGCGCTGGTGACATATCTCTGAGCTGGGATGACATCTGCTACTCAAGG	962
Db	62	TCCAGGCGCTGGTGACATATCTCTGAGCTGGGATGACATCTGCTACTCAAGG	121
QY	963	AGCCCTGTGCTTCAGAGGCTGGCGGCTCCTGTGCAAGGATATACCCCTACCTGTGA	1022
Db	122	AGCCCTGTGCTTCAGAGGCTGGCGGCTCCTGTGCAAGGATATACCCCTACCTGTGA	181
QY	1023	CTGTGCAAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCCTGTTTCTGAAG	1082
Db	182	CTGTGCAAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCCTGTTTCTGAAG	241
QY	1083	CTGCCACAGAGGAGCTCTTCTCAGTGAGGCTTCCGGAGTCCCTCAGCTTCTACA	1142
Db	242	CTGCCACAGAGGAGGAGTCTTCTCAGTGAGGCTTCCGGAGTCCCTCAGCTTCTACA	301
QY	1143	TCAGCCTGAATGAGGAGCTCTCTTTGATGATGCCTAG	1183
Db	302	TCAGCCTGAATGAGGAGCTCTCTTTGATGATGCCTAG	342

## RESULT 6

US-09-814-353-17314  
; Sequence 17314, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-0065

; CURRENT APPLICATION NUMBER: US/09/814,353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17314  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-814-353-17314

Query Match 18.0%; Score 213.4; DB 10; Length 320;  
Best Local Similarity 95.2%; Pred. No. 5.5e-55;  
Matches 220; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY	536	GGCCCGCGCGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGA	595
Db	90	GTCCGCGCGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGA	149
QY	596	GACTGTGAGCGTCTCTCTGAAGTCTCAGGAGAGATATAACATCCCGAGGTCCAC	655
Db	150	GACTGTGAGCGTCTCTCTGAAGTCTCAGGAGAGATATAACATCCCGAGGTCCAC	209
QY	656	GTGGGCAAGTCTCCCATGGGTGCTGATGAGGCTGACGAGGAGAAAGCAGAGAA	715
Db	210	GTGGGCAAGTCTCCCATGGGTGCTGATGAGGCTGACGAGGAGAAAGCAGAGAA	269
QY	716	CTGCTGTTGTTTACCTGGGAACCTCGAGGGGCTTCTCTCATCCGGGAGAGC	766
Db	270	CTGCTGTTGTTTACCTGGGAACCTCGAGGGGCTTCTCTCATCCGGGAGAGC	320

## RESULT 7

US-09-954-456-499  
; Sequence 499, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc

; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863

; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 499  
; LENGTH: 2665  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-499

Query Match 13.3%; Score 157.4; DB 9; Length 2665;  
Best Local Similarity 54.2%; Pred. No. 1.2e-37;  
Matches 354; Conservative 0; Mismatches 281; Indels 18; Gaps 1;  
QY 410 CCAGAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCTCAAGGCCAGGGA 469  
Db 24 CCAGGAAAAAGAAATGGGAAACAGCATGAATCCACCCCTGGCGCTCCCGAGAGG 83  
QY 470 CCTGTGACCATGAAGCAGAGAGAGCAAGGCCACAGCCGTCGGCCCTGGGCAAGTTCCCG 529  
Db 84 CCCTGCCCAACCCGAGGAGCTGGATAGCGACTTCCTTGGCGTGTCTAAGTGACTACCCG 143  
QY 530 GCAGGTGGCCCGCCGAGCTGTCTGTGAGACTCGGGAGCCATTGACCATCTCTCTGAG 589  
Db 144 TCTCTTGACATCAGCCCCCGATATTTCCGCCGAGGGGAGAACTGGTGTGATTTCTGAT 203  
QY 590 GATGGAGACTGTGGACGCTGTCTGTGAAGTCTCAGGCGAGAGATATAACATCCCCAGC 649  
Db 204 GAAGGGGCTGTGGAAAGCTATTCTCTTAGCACTGGTTCGAGAGATTATATCCCTCGA 263  
QY 650 GTCCAGCTGGGCAAAAGTCTCCATGGGTGGCTGTATGAGGGCTCAGCAGGAGGAGAAACA 709  
Db 264 ATATGTGTGGCCAGAGTTTACCATGGCTGGCTGTGGGGCTGGGCGAGAGCAAGGCC 323  
QY 710 GAGGAACTCTGTGTTACTCTGTGTCAGTCGCGCTCAGCGCCCTGATCTCCGGAGCCAG 769  
Db 324 GAGGAGCTGCTCAGCTGCGACACAAAGTTCGGCTCTTCAATGATCAGAGAGAGTGAG 383  
QY 770 ACCAGGAGAGGCTTACTCTGTCTGTGTCAGTCGCGCTCAGCGCCCTGATCTCCGGAGCC 829  
Db 384 ACCAAGAAAGGTTTACTCATCTGCTGTGAGACAGGCA-----G 425  
QY 830 ATCAGACTACAGGATCAGTCGCTTGACATGCGCTGTGATCATCTCACCCTCGCTC 889  
Db 426 GTAAAGCAATTACCGATTTTCGCTGTGCGCAAACTGGTACTCATTTCCCGAGGCTC 485  
QY 890 ACCTTCCCTCACTCAGGCTGTGTCAGGAGGCTGCGCCCTCCCTGGCAAGGATATA 1009  
Db 486 ACCTTCCAGTGCCTGGAGACTGTGTGAACCACTATTCTGAGGTGGCTGTGATGGCTGTGC 545  
QY 950 TGCCTACTCAAGGAGCCCTGTGTCTGTGAGAGGCTGCGCCCTCCCTGGCAAGGATATA 1009  
Db 546 TGTGTGCTCACCACGCCCTGCTGACACAAAGCAGCGCTGCCAGCAGTGAGGGCTCC 605  
QY 1010 CCCCTACCTGTGACTGTGCGAGGAGCACCCTCACTGGAAGAGCTGGACAG 1062  
Db 606 AGCTCACCTGTACCTTGGTCAGAGAGCTGTGAGCTGGAGAGAGTGTCAG 658

RESULT 8  
US-10-002-600-91  
; Sequence 91, Application US/10002600  
; Publication No. US20020137077A1  
; GENERAL INFORMATION:  
; APPLICANT: Hopkins, Christopher M.  
; APPLICANT: Peterson, David P.  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Hawkins, Phillip R.  
; TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS  
; FILE REFERENCE: PA-0042 US  
; CURRENT APPLICATION NUMBER: US/10/002,600  
; PRIOR FILING DATE: 2001-10-25  
; PRIOR APPLICATION NUMBER: 60/243,521  
; PRIOR FILING DATE: 2000-10-25

; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PERL Program  
; SEQ ID NO 91  
; LENGTH: 3756  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Template ID: 059263.15  
US-10-002-600-91

Query Match 13.3%; Score 157.4; DB 13; Length 3756;  
Best Local Similarity 54.2%; Pred. No. 1.4e-37;  
Matches 354; Conservative 0; Mismatches 281; Indels 18; Gaps 1;  
QY 410 CCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCTCAAGGCCAGGGA 469  
Db 1098 CCAGGAAAAAGAAATGGGAAACAGCATGAATCCACCCCTGGCGCTCCCGAGAGG 1157  
QY 470 CCTGTGACCATGAAGCAGAGAGAGCAAGGCCACAGCCGTCGGCCCTGGGCAAGTTCCCG 529  
Db 1158 CCCTGCCCAACCCGAGGAGCTGGATAGCGACTTCCTTGGCGTGTCTAAGTGACTACCCG 1217  
QY 530 GCAGGTGGCCCGCCGAGCTGTCTGTGAGACTCGGGAGCCATTGACCATCTCTCTGAG 589  
Db 1218 TCTCTGACATCAGCCCCCGATATTTCCGCCGAGGGGAGAACTGGCTGTGATTTCTGAT 1277  
QY 590 GATGGAGACTGTGGACGCTGTCTGTGAAGTCTCAGGCGAGAGATATAACATCCCCAGC 649  
Db 1278 GAAGGGGCTGTGGAAAGCTATTTCTCTTAGCACTGGTTCGAGAGAGTTTACATCCCTGGA 1337  
QY 650 GTCCAGCTGGGCAAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGGAGAAAGCA 709  
Db 1338 ATATGTGTGGCCAGAGTTTACCATGCTGGCTGTGTTGAGGGCTGGGCGAGACAAGGCC 1397  
QY 710 GAGGAACTCTGTGTTTACTCTGTGTCAGTCGCGCTCAGGAGGCGCTTCTCATCCGGAGAGCCAG 769  
Db 1398 GAGGAGCTGTGTCAGCTGCGCAGACAAAGGTCGGCTCTTCAATGATCAGAGAGAGTGAG 1457  
QY 770 ACCAGGAGAGGCTTACTCTGTCTGTCAGTCGCGCTCAGCGCCCTGATCTCCCTGGAGCCCG 829  
Db 1458 ACCAAGAAAGGTTTACTCATCTGCTGTGAGCAAAAGCA-----G 1499  
QY 830 ATCAGACTACAGGATCAGTCGCTTGACATGCGCTGTGATCATCTCACCCTCGCTC 889  
Db 1500 GTAAAGCAATTACCGATTTTCCGCTGTGCGCAAACTGGTACTCATATTCCCGAGGCTC 1559  
QY 890 ACCTTCCCTCACTCAGGCTGTGTCAGGAGGCTGTGTGAACCACTATTCTGAGGTGGCTGTGTC 949  
Db 1560 ACCTTCCAGTGCCTGGAGACTGTGTGAACCACTATTCTGAGGTGGCTGTGTC 1619  
QY 950 TGCCTACTCAAGGAGCCCTGTGTCTGTGAGAGGCTGCGCCCTCCCTGGCAAGGATATA 1009  
Db 1620 TGTGTGCTCACCACGCCCTGCTGACACAAAGCAGCGCTGCCAGCAGTGAGGGCTCC 1679  
QY 1010 CCCCTACCTGTGACTGTGCGAGGAGCACCCTCACTGGAAGAGCTGGACAG 1062  
Db 1680 AGCTCACCTGTCTACCTTGGCTCAGAAGACTGTGGAGCTGGAGAGAGTGTCAG 1732

RESULT 9  
US-09-864-761-2829  
; Sequence 2829, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemica-X-1  
; FILE REFERENCE: Aemica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761

```

; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2829
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031662.24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
US-09-864-761-2829

Query Match 12.0%; Score 141.8; DB 9; Length 432;
Best Local Similarity 95.4%; Pred. No. 4.7e-33;
Matches 146; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 912 TGGTGGACCAATTACTCTGAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTG 971
Db 253 TGGAGGTCTCTTCTCTAGAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTG 312
QY 972 TCCTGCGAGGGGTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGA 1031
Db 313 TCCTGCGAGGGGTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGA 372
QY 1032 GGACACCACTCAACTGGAAGAGCTGCAGACT 1064
Db 373 GGACACCACTCAACTGGAAGAGCTGCAGAGGT 405

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15513
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031662.24
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
US-09-864-761-15513

Query Match 12.0%; Score 141.8; DB 9; Length 448;
Best Local Similarity 95.4%; Pred. No. 4.7e-33;
Matches 146; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 912 TGGTGGACCAATTACTCTGAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTG 971
Db 269 TGGAGGTCTCTTCTCTAGAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTG 328
QY 972 TCCTGCGAGGGGTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGA 1031

```

Db 329 TCCTGAGAGGGCTGCCCGCTCCCTGGCAGGATATACCTCTACCTGTGACTGTGCAGA 388  
QY 1032 GGACACCACTCAACTGGAAGAGCTGGACAGCT 1064  
Db 389 GGACACCACTCAACTGGAAGAGCTGGACAGGT 421

## RESULT 11

US-09-814-353-4631  
; Sequence 4631, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4631  
; LENGTH: 152  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; LOCATION: 17, 102, 112  
; OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-4631

Query Match 11.1%; Score 131.6; DB 10; Length 152;  
Best Local Similarity 95.0%; Pred. No. 4.8e-30;  
Matches 134; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 536 GGCCCGGCGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGA 595  
Db 12 GTCGNGCCGAGGTGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGA 71  
QY 596 GACTGTGGAGCGGTGCTGCTGAGTCTCAGGAGAGATATACATCCCCAGCGTCCAC 655  
Db 72 GACTGTGGAGCGGTGCTGCTGAGTCTCANGCAGAGATNTAACATCCCCAGCGTCCAC 131  
QY 656 GTGGCCAAAGTCTCCCATGGG 676  
Db 132 GTGGCCAAAGTCTCCCATGGG 152

## RESULT 12

US-09-814-353-10930  
; Sequence 10930, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; THERAPY OF OVARIAN CANCER

; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10930  
; LENGTH: 152  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; LOCATION: 17, 102, 112  
; OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-10930

Query Match 11.1%; Score 131.6; DB 10; Length 152;  
Best Local Similarity 95.0%; Pred. No. 4.8e-30;  
Matches 134; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 536 GGCCCGGCGAGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGA 595  
Db 12 GTCGNGCCGAGGTGCTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGA 71  
QY 596 GACTGTGGAGCGGTGCTGCTGAGTCTCAGGAGAGATATACATCCCCAGCGTCCAC 655  
Db 72 GACTGTGGAGCGGTGCTGCTGAGTCTCANGCAGAGATNTAACATCCCCAGCGTCCAC 131  
QY 656 GTGGCCAAAGTCTCCCATGGG 676  
Db 132 GTGGCCAAAGTCTCCCATGGG 152

## RESULT 13

US-10-062-674-2038  
; Sequence 2038, Application US/10062674  
; Publication No. US20040005559A1  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.  
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
; FILE REFERENCE: PA-0028-1 CIP  
; CURRENT APPLICATION NUMBER: US/10/062,674  
; CURRENT FILING DATE: 2002-01-30  
; PRIOR APPLICATION NUMBER: US 09/625,102  
; PRIOR FILING DATE: 2000-07-24  
; NUMBER OF SEQ ID NOS: 2217  
; SOFTWARE: PEEL Program  
; SEQ ID NO 2038  
; LENGTH: 2343  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040005559A1 411296.5  
US-10-062-674-2038

Query Match 10.5%; Score 124.6; DB 15; Length 2343;  
Best Local Similarity 54.1%; Pred. No. 1.3e-27;  
Matches 284; Conservative 0; Mismatches 229; Indels 12; Gaps 1;  
QY 457 CCAAGCCAGGACCTGTGACCATGGAGCAGAGAGCAAGCCACGCCCTGCGCCT 516

Db	609	CAACAGCAACACA	CACAGGAATCAGGAGGCAGGCTCTGAGGACATCATCGTGGTGCCT	668	
Qy	517	GGGCAGTTTCCCGG	CAGGTGGCCGGCCGAGCTGCGTGTAGACTCGGGAGCCATTGCAC	576	
Db	669	GTATGATTACGAGGCCA	TTCACCAAGAGACTCAGCTTCAGAAGGGGACCAGATGGT	728	
Qy	577	CATCGTCTCTGAGGATG	GAGACTGCTGTGACGGTGTCTGTAAGTCTCAGGCGAGAGTA	636	
Db	729	GGTCTTAGAGGAATCCG	GGGAGTGGTGAAGGCTCGATCCCTGGCCACCCGGAAGGAGGG	788	
Qy	637	TAACATCCCGACGGTCC	ACGTGGGCAAAAGTCTCCCAT-----GGTGGTCTCTGA	684	
Db	789	CTACATCCCAAGCAACT	ATGATCGCCCGGTTGACTCTCTGGAGACAGAGAGTGGTTTTT	848	
Qy	685	TGAGGGCTGTAGCAGG	GAGAAAGACAGAGAACTGCTGTTGTTA	CTGTGGAAACCTGTGAGG	744
Db	849	CAAGGGCATAGCGGGA	AGGACGACAGCGCCCACTGCTGGCTCCCGCAACATGCTGGG	908	
Qy	745	GGCCTTCCTCATCCGGG	AGAGCCAGACCAGAGAGGCTTTACTCTCTGTGTCAGTCCGCT	804	
Db	909	CTCCTTCATGATCCGGG	ATACGGAGACCACTTAAGGAGCTACTCTCTGTCCGTGGCAGA	968	
Qy	805	CAGCGGCCCTGCATCT	GTGGGACCGGATCAGACATTA	CAGGATCCACTGCCTTGACAATGG	864
Db	969	CTACGACCCCTCGG	CAGGAGATACCGTGAACATTACAAGATCCGGACCCCTGGACAACGG	1028	
Qy	865	CTGCGCTGTACATCT	CACCGCCCTCACCTTCCCTCTCACTCCAGAGCCCTGTGGACCACTTA	924	
Db	1029	GGGCTTCTACATAT	CCCCCGAAGCACTTCAGCACTCTGAGGAGCTGTGGGACCACCTA	1088	
Qy	925	CTCTGAGCTGGCGAT	GCATCTGCTGCCTACTCAAGAGAGCCCTG	969	
Db	1089	CAAGAGGGGAAACAG	CGGCTCTGCCAGAACTGTGCTGTGCCTTG	1133	

RESULT 14  
US-10-193-720-1  
; Sequence 1, Application US/10193720  
; Publication No. US20040009173A1  
; GENERAL INFORMATION:  
; APPLICANT: Priera, Annabelle M.  
; APPLICANT: Wong, Brian R.  
; APPLICANT: Masuda, Estaban  
; APPLICANT: Powell, Mark  
; TITLE OF INVENTION: Modulators of Leukocyte Activation, HCK Compositions and Methods  
; FILE REFERENCE: A-71313/RMS/DHR  
; CURRENT APPLICATION NUMBER: US/10/193,720  
; CURRENT FILING DATE: 2002-11-27  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1924  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (78)..(1595)  
; OTHER INFORMATION:  
; US-10-193-720-1

332	GGTCTCTAGAGGAATCGGGGAGTGTGAAGGCTCGATCCCTGGCCACCCGGAAGAGGG	391
367	TAACATCCCACGGGTCACGTGGGCAAAAGTCTCCCAT-----GGGTGGCTGTA	426
392	CTACATCCCACGAACAATATGTCCCGGGTTGACTCTCTGGAGACAGAGGAGTGGTTTT	451
417	TGAGGSCCTGACGACGGAGGAAGCAGAGGAACACTGCTGTGTACCTGGGAACCTCGAGG	482
442	CAAGGSCATACGCGGGAAGGACGAGGCGCAACTGCTGGTCCCGCAACATGCTGGG	511
467	GGCTTTCCTCATCCGGGAGACGACAGCAGGAGGCTCTTACTCTCTGTCACTCCGCCT	532
492	CTCTTTTCATGATCCGGGATACGAGACCACTAAAGGAAGCTACTCTTTGTCCGTGGAGA	561
517	CAGCGCCCTTGATCCTGGGACCGGATCAGACACTACAGGATCCACTGCTCTGACAAATGG	586
542	CTACGACCTCGGACGGGAGATACCGTGAACATTACAGATCGGACCTGGACAAACGG	611
567	CTGGCTGTACATCTCACCGCGCTCACCTTTCCTTACTCCAGGCGCTGGTGAACATTA	632
592	GGGCTTCTACATATCCCCCGGAACAGCTTCAGCACTCTCGAGGAGCTGTGGACCACTA	661
617	CTCTGAGTGGCGATGACATCTGCTGCTTACTCAAGGAGCCCTG	686
642	CAAGAAGGGGAACGACGGGCTCTCCGAGAACCTCGTGCCCTG	716

RESULT 15  
US-09-954-456-1983  
Sequence 1983, Application US/09954456  
Patent No. US20020115057A1  
GENERAL INFORMATION:  
APPLICANT: Young, Paul  
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents  
TITLE OF INVENTION: Sets  
FILE REFERENCE: 689290-76  
CURRENT APPLICATION NUMBER: US/09/954,456  
CURRENT FILING DATE: 2001-09-18  
PRIOR APPLICATION NUMBER: US/60/233,617  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/234,052  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234,923  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,134  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,637  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,638  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,711  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,720  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,840  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,863  
PRIOR FILING DATE: 2000-09-27  
NUMBER OF SEQ ID NOS: 2276  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1983  
LENGTH: 2015  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-954-456-1983

Query Match 10.4%; Score 123; DB 9; Length 2015;  
 Best Local Similarity 53.9%; Pred. No. 3.9e-27;  
 Matches 283; Conservative 0; Mismatches 230; Indels 12; Gaps 1;  
 QY 457 CCAGGCCAGGACCTGTGCACATGGACGACGAGAGAGCAAGGCCACAGCCGTGGCCCT 516

```
Db 303 CAACAGCAACACACAGGAATCAGGAGGAGGCTCTGAGGACATCATCGTGGTTGCCCT 362
QY 517 GGGCAGTTTCCGGCAGGTGGCCCGCCAGCTGTGCTGAGACTCGGGAGCCATTGAC 576
Db 363 GTATGATTAAGAGGCCATTCCACAGAGACCTCAGCTTCCAGAAGGGGGACCCAGATGGT 422
QY 577 CATCGTCTCTGAGGATGGAGACTGGGTGACGGTGTCTGTGAAGTCTCAGGCAGAGAGTA 636
Db 423 GGTCTCTAGAGGAATCCGGGGAGTGGTGAAGCTCGATCCCTGGCCACCCGGAGAGGGG 482
QY 637 TAACATCCCCAGGGTCCAGCTGGGCAAGTCTCCCAT-----GGGTGGCTGTA 684
Db 483 CTACATCCCAAGCAACTATGTGCGCGGGTGAAGTCTCTGGAGACAGAGGAGTGGTTTTT 542
QY 685 TCAGGGCCTGAGCAGGGAGAAAGCAGAGAACTGTGTGTGTTTACCTGGGAACCCCTGGAGG 744
Db 543 CAAGGGCATCAGCCGGAAGGACGACAGAGGCCCACTGTGGCTCCCGCAACATGCTGGG 602
QY 745 GGCCTTCCTCATCCGGGAGAGCCACAGAGGAGGCTTTACTCTGTCTGTGAGTCCGCT 804
Db 603 CTCCTTCATGATCCGGGATAGCGAGACCACCTAAAGGAAGCTACTCTTTGTCCGTGCGAGA 662
QY 805 CAGCCGCCCTGCATCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTGACAAATGG 864
Db 663 CTACGACCTCGGAGGAGATACCGTGAACATTACAAGATCCGACCCCTGGACAACGG 722
QY 865 CTGGCTGTACATCTACCGCGCTCACCTTCCCTCACTCCAGGCCCTGGTGGACCATTA 924
Db 723 GGGCTTCTACATATCCCCCGAAGCACCTTCAGCACTCTGCAGGAGCTGGTGGACCACTA 782
QY 925 CTCCTGAGCTGGGGATGACATCTGCTGCTCTACTCAAGGAGCCCTG 969
Db 783 CAAGAAGGGAGACGACGGGCTCTGCCAGAAACTGTGCGTGCCCTG 827
```

Search completed: February 20, 2004, 05:59:08  
Job time : 2861.55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 01:45:20 ; Search time 4572 Seconds  
(without alignments)  
11214.963 Million cell updates/sec

Title: US-09-939-853A-74  
Perfect score: 1183  
Sequence: 1 agctagagctccaaggaccc.....tctcttgatgacgctag 1183

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.man.\*
- 37: em.htg.vrt.\*
- 38: em.ly.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183	100.0	1183	6	AX443133 Sequence
2	1183	100.0	1183	6	AX443135 Sequence
3	1132	95.7	2567	6	AX452880 Sequence
4	1046	88.4	2538	9	BC042041 Homo sapi
5	735	62.1	786	6	AX511153 Sequence
6	735	62.1	786	6	AX512845 Sequence
7	735	62.1	786	9	AF290985 Homo sapi
8	735	62.1	786	9	AF290985 Homo sapi
9	735	62.1	786	9	AF290985 Homo sapi
10	684	57.8	2788	6	AX780857 Sequence
11	483	40.8	737	6	AX511155 Sequence
12	483	40.8	737	9	AP290986 Homo sapi
13	354	29.9	66741	9	HS460J8
14	354	29.9	145833	2	AC026539 Homo sapi
15	152	12.8	145068	9	HS460J8
16	152	12.8	145833	2	AC026539 Homo sapi
17	43	3.6	777	6	AX511151 Sequence
18	43	3.6	1321	10	AF434390 Mus muscu
19	43	3.6	1348	6	AX511150 Sequence
20	43	3.6	1384	10	BC052655 Mus muscu
21	43	3.6	1384	10	BC052655 Mus muscu
22	41	3.5	163240	10	AL935150 Mouse DNA
23	30	2.5	30	6	AX452884 Sequence
24	29	2.5	299947	2	AC125701 Rattus no
25	29	2.5	322940	2	AC123560 Rattus no
26	26	2.2	26	6	AX443200 Sequence
27	25	2.1	25	6	AX511164 Sequence
28	25	2.1	25	6	AX511165 Sequence
29	25	2.1	25	6	AX511166 Sequence
30	23	1.9	23	6	AX511163 Sequence
31	23	1.9	8028	2	AC014511 Drosophil
32	23	1.9	14560	10	AB011527 Rattus no
33	23	1.9	177577	3	AC022349 Drosophil
34	23	1.9	179016	3	AC018489 Drosophil
35	23	1.9	189516	2	AC128065 Rattus no
36	23	1.9	270729	2	AC133702 Rattus no
37	23	1.9	300469	3	AE003503 Drosophil
38	22	1.9	22	6	AX443201 Sequence
39	22	1.9	475	9	HSPPROT04
40	22	1.9	3070	6	E24027 Human.P pro
41	22	1.9	3070	6	E24028 Method for
42	22	1.9	3070	9	M99564 Homo sapien
43	22	1.9	3082	9	BC012097 Homo sapi
44	22	1.9	67718	2	AC135349 Homo sapi
45	22	1.9	67718	2	AC135349 Homo sapi

ALIGNMENTS

RESULT 1	AX443133	Sequence 74 from Patent WO0216599.	1183 bp	DNA	linear	PAT 02-JUL-2002
AX443133	Sequence 74 from Patent WO0216599.					
LOCUS	AX443133					
DEFINITION	AX443133					
ACCESSION	AX443133.1	GI:21690555				
VERSION	AX443133.1					
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,  
Shmukets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,  
Topper,J.N. and Yang,R.B.

TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES Location/Qualifiers  
source 1..1183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1183; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGCTAGAGCTCAAGACACCCAGCGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
DB 1 AGCTAGAGCTCAAGACACCCAGCGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
QY 61 CTTTCCTCCCTGGCTCGGCTCGGCTGTGTGAGAGGTTCCCAAGTCCAGATCCCTAAGAG 120  
DB 61 CTTTCCTCCCTGGCTCGGCTCGGCTGTGTGAGAGGTTCCCAAGTCCAGATCCCTAAGAG 120  
QY 121 CATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180  
DB 121 CATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180  
QY 181 ACCCAACACACCTAGCTCTCCCTGAAGATCTCCAGAGTCCAGAGTGTCTGGGTG 240  
DB 181 ACCCAACACACCTAGCTCTCCCTGAAGATCTCCAGAGTGTCTGGGTG 240  
QY 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTGCCA 300  
DB 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTGCCA 300  
QY 301 GCCAGAGATGGTCTCAGAGAGTGTCTTCCCAAGCTTTGATGACAAACCAATTTC 360  
DB 301 GCCAGAGATGGTCTCAGAGAGTGTCTTCCCAAGCTTTGATGACAAACCAATTTC 360  
QY 361 CTCGATGATGTCTCTCAGTGTCTCTGAGGAGCAATGGAGTCTGCCAGAGAG 420  
DB 361 CTCGATGATGTCTCTCAGTGTCTCTGAGGAGCAATGGAGTCTGCCAGAGAG 420  
QY 421 AAAATCTGCGCAAGCCCAAGCTTGAAGTTCCTGTCTCAAGGGCAGGAGCTGTGACCAT 480  
DB 421 AAAATCTGCGCAAGCCCAAGCTTGAAGTTCCTGTCTCAAGGGCAGGAGCTGTGACCAT 480  
QY 481 GGAAGCAGAGAGAGAGGACAGCCGCTGGCCCTGGCAGTTCCTGGCAGGTGGCC 540  
DB 481 GGAAGCAGAGAGAGAGGACAGCCGCTGGCCCTGGCAGTTCCTGGCAGGTGGCC 540  
QY 541 GGCGAGCTGTGCTGAGATCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGCTG 600  
DB 541 GGCGAGCTGTGCTGAGATCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGCTG 600  
QY 601 GTGACAGTGTCTGAGTCTCAGCAGAGAGTATAACATCCCAAGGTCCAGGTGG 660  
DB 601 GTGACAGTGTCTGAGTCTCAGCAGAGAGTATAACATCCCAAGGTCCAGGTGG 660  
QY 661 CAAAGTCTCCCAAGGGTGTATGAGGCGCTGAGCAGGAGAGAAAGCAGAGAACTGCT 720  
DB 661 CAAAGTCTCCCAAGGGTGTATGAGGCGCTGAGCAGGAGAGAAAGCAGAGAACTGCT 720  
QY 721 GTTCTTACCTGGGAACCTGGAGGGCTTCTCCTATCCCGGAGAGCCAGACAGGAGG 780  
DB 721 GTTCTTACCTGGGAACCTGGAGGGCTTCTCCTATCCCGGAGAGCCAGACAGGAGG 780  
QY 781 CTCTTACTCTGTGAGTCCGCTCAGCCGCTTGCATCTCTGGGACCGGATCAGACACTA 840  
DB 781 CTCTTACTCTGTGAGTCCGCTCAGCCGCTTGCATCTCTGGGACCGGATCAGACACTA 840  
QY 841 CAGATCCATGCTTGAATGGTGGTGTGATCTTACCGGCTTCACTTCCCTC 900  
DB 841 CAGATCCATGCTTGAATGGTGGTGTGATCTTACCGGCTTCACTTCCCTC 900

QY 901 ACTCCAGGCTGGTGGACCACTTACTCTGAGCTGGGGATGACATCTGCTGCTACTCAA 960  
DB 901 ACTCCAGGCTGGTGGACCACTTACTCTGAGCTGGGGATGACATCTGCTGCTACTCAA 960  
QY 961 GGAGCCCTGTCTCTGACAGAGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACTGT 1020  
DB 961 GGAGCCCTGTCTCTGACAGAGGCTGGCCGCTCCCTGGCAAGGATATACCCCTACTGT 1020  
QY 1021 GACTGTGACAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTTCTGA 1080  
DB 1021 GACTGTGACAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTTTTCTGA 1080  
QY 1081 AGCTGCCACAGGGAGAGTCTCTTCTCAGTGAAGGTCTCCGGAGTCCCTCAGCTCTA 1140  
DB 1081 AGCTGCCACAGGGAGAGTCTCTTCTCAGTGAAGGTCTCCGGAGTCCCTCAGCTCTA 1140  
QY 1141 CATCAGCCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183  
DB 1141 CATCAGCCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183  
RESULT 2  
AX443135/c  
LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002  
DEFINITION Sequence 76 from Patent WO0216599.  
ACCESSION AX443135  
VERSION AX443135.1 GI:21690556  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,  
Shinkens,R.A., Sytek,K.A., Szekeres,E.S., Tomlinson,J.E.,  
Topper,J.N. and Yang,R.B.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES  
1..1183  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1183; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGCTAGAGCTCAAGACACCCAGCGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGG 60  
DB 1183 AGCTAGAGCTCAAGACACCCAGCGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGG 1124  
QY 61 CTTTCCTCCCTGGCTCGGCTGTGTGGAGGTTCCCAAGTCCAGATCCCTAAGAG 120  
DB 1123 CTTTCCTCCCTGGCTCGGCTGTGTGGAGGTTCCCAAGTCCAGATCCCTAAGAG 1064  
QY 121 CATGGGAGCTGATTCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 180  
DB 1063 CATGGGAGCTGATTCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTGAGCT 1004  
QY 181 ACCCAACCAACACTAGCTCTCCCTGAGATCCTCCAGGCTGAGAGATCTGGGTG 240  
DB 1003 ACCCAACCAACACTAGCTCTCCCTGAGATCCTCCAGGCTGAGAGATCTGGGTG 944  
QY 241 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTGCCA 300  
DB 943 TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGCCCCCAAGCCCTAAGCTGCCA 884  
QY 301 GCCAGAGATGCTCTCAGCAGAGCTGTCTTCCCAAGCTTTGATGACAAACCAATTTC 360

883 GCCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAATTTCC 824  
 361 CTCGATGATGTCTTCTGAGTGTCTGCTGAGGAAACAATGGGAGTCTCCAGCAGAGAG 420  
 823 CTCGATGATGTCTTCTGAGTGTCTGCTGAGGAAACAATGGGAGTCTCCAGCAGAGAG 764  
 421 AAAATCTCTGCCAAGCCCAAGCTGTAGTTCTCTGTGCTCCAAAGCCAGGAGCTGTGACCAT 480  
 763 AAAATCTCTGCCAAGCCCAAGCTGTAGTTCTCTGTGCTCCAAAGCCAGGAGCTGTGACCAT 704  
 481 GGAAGCAGAGAGAGCAAGCCCAAGCTGTAGTTCTCTGTGCTCCAAAGCCAGGAGCTGTGACCAT 540  
 703 GGAAGCAGAGAGAGCAAGCCCAAGCTGTAGTTCTCTGTGCTCCAAAGCCAGGAGCTGTGACCAT 644  
 541 GGCAGAGTGTCTGCTGAGACTCGGGAGCCATGACCATGCTCTGTGAGGATGGAGACTG 600  
 643 GGCAGAGTGTCTGCTGAGACTCGGGAGCCATGACCATGCTCTGTGAGGATGGAGACTG 584  
 601 GTGAGCGGTGTCTGCTGAGACTCGGGAGCCATGACCATGCTCTGTGAGGATGGAGACTG 660  
 583 GTGAGCGGTGTCTGCTGAGACTCGGGAGCCATGACCATGCTCTGTGAGGATGGAGACTG 524  
 661 CAAAGTCTCCATGGGTGCTGTATGAGGGCTGTGAGGGAGGAGAAAGCAGAGGAACTGCT 720  
 523 CAAAGTCTCCATGGGTGCTGTATGAGGGCTGTGAGGGAGGAGAAAGCAGAGGAACTGCT 464  
 721 GTTGTACTGTGGAACTCTGAGGAGGCTTCTCTATCCGGAGAGCCAGAGAGAGG 780  
 463 GTTGTACTGTGGAACTCTGAGGAGGCTTCTCTATCCGGAGAGCCAGAGAGAGG 404  
 781 CTCTTACTCTGTGCTGAGTCCGCTCAGCGGCTGTGAGGGAGGAGAAAGCAGAGGAACTGCT 840  
 403 CTCTTACTCTGTGCTGAGTCCGCTCAGCGGCTGTGAGGGAGGAGAAAGCAGAGGAACTGCT 344  
 841 CAGGATCACTGCTGTGACATGGGTGCTGTATCTACCGGCTCTACCTTCCGCTC 900  
 343 CAGGATCACTGCTGTGACATGGGTGCTGTATCTACCGGCTCTACCTTCCGCTC 284  
 901 ACTCAGGCTCTGTGAGTCCGCTCAGCGGCTGTGAGGGAGGAGAAAGCAGAGGAACTGCT 960  
 283 ACTCAGGCTCTGTGAGTCCGCTCAGCGGCTGTGAGGGAGGAGAAAGCAGAGGAACTGCT 224  
 961 GGAGCCCTGTCTGTGAGAGGCTGTGAGGCTGTGAGGAGGAGAAAGCAGAGGAACTGCT 1020  
 223 GGAGCCCTGTCTGTGAGAGGCTGTGAGGCTGTGAGGAGGAGAAAGCAGAGGAACTGCT 164  
 1021 GACTGTGAGAGGAGCAACCACTCACTGAGAGAGCTGTGAGGAGGAGAAAGCAGAGGAACTGCT 1080  
 163 GACTGTGAGAGGAGCAACCACTCACTGAGAGAGCTGTGAGGAGGAGAAAGCAGAGGAACTGCT 104  
 1081 AGCTGCCACAGGAGGAGTCTTCTCAGTGGGCTGTGAGGAGGAGAAAGCAGAGGAACTGCT 1140  
 103 AGCTGCCACAGGAGGAGTCTTCTCAGTGGGCTGTGAGGAGGAGAAAGCAGAGGAACTGCT 44  
 1141 CATCAGCTGATCAGGAGGAGTCTTCTCAGTGGGCTGTGAGGAGGAGAAAGCAGAGGAACTGCT 1  
 43 CATCAGCTGATCAGGAGGAGTCTTCTCAGTGGGCTGTGAGGAGGAGAAAGCAGAGGAACTGCT 1

RESULT 3  
 AX452880  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villar, J.J. and  
 Kanner, S.B.

TITLE Cloning and expression of human slap-2: a novel sh2/sh3 domain-containing human slap homologue having immune cell-specific expression  
 JOURNAL Patent: WO 0242457-A 1 30-MAY-2002; Bristol-Myers Squibb Co. (US)  
 FEATURES  
 source 1..2567  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

Query Match 95.7%; Score 1132; DB 6; Length 2567;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGCTAGAGCTCCAAAGGAGCCAGCCCTGTCTGTCTGTGACAGAGCTCAAGAGGCGCTGGG 60  
 Db 18 AGCTAGAGCTCCAAAGGAGCCAGCCCTGTCTGTCTGTGACAGAGCTCAAGAGGCGCTGGG 77  
 Qy 61 CCTTCCCTCCCTCGCTGGCTGTCTGTGGAGGGTTCCTCAGTCCAGATCCCTAAGGAG 120  
 Db 78 CCTTCCCTCCCTCGCTGGCTGTCTGTGGAGGGTTCCTCAGTCCAGATCCCTAAGGAG 137  
 Qy 121 CATGGGCGAGCTGATCCATCCCTGTGTGTACAACTGCTCACTGCACAGACAGATGCTGAGCT 180  
 Db 138 CATGGGCGAGCTGATCCATCCCTGTGTGTACAACTGCTCACTGCACAGACAGATGCTGAGCT 197  
 Qy 181 ACCCAACCAACACCTAGAGCTCTCCCTGAGAGTCTCCAGGCTGAGAGTCTGGGTG 240  
 Db 198 ACCCAACCAACACCTAGAGCTCTCCCTGAGAGTCTCCAGGCTGAGAGTCTGGGTG 257  
 Qy 241 TCCTAGGACCAAGGACACTGGCAGAGCTTCCAGAGGGGCCCCAAAGCCCTTAACCTGTCCA 300  
 Db 258 TCCTAGGACCAAGGACACTGGCAGAGCTTCCAGAGGGGCCCCAAAGCCCTTAACCTGTCCA 317  
 Qy 301 GCCAGAGCATGTCTCAGCAGAGCTGTCTTCCAGAGCTTTGATGACAAACCAATTTCC 360  
 Db 318 GCCAGAGCATGTCTCAGCAGAGCTGTCTTCCAGAGCTTTGATGACAAACCAATTTCC 377  
 Qy 361 CTCGATGATGTCTTCTGAGTGTCTCTGAGTGAAGAACATGGGAAGTCTGCCAGCAGAG 420  
 Db 378 CTCGATGATGTCTTCTGAGTGTCTCTGAGTGAAGAACATGGGAAGTCTGCCAGCAGAG 437  
 Qy 421 AAAATCTCTGCCAAGCCCAAGCTGTGAGTTCCTGTCCAAGGCCAGGAGCTGTGACCAT 480  
 Db 438 AAAATCTCTGCCAAGCCCAAGCTGTGAGTTCCTGTCCAAGGCCAGGAGCTGTGACCAT 497  
 Qy 481 GGAAGCAGAGAGAGCAAGCCCAAGCCGACAGCCGTGGCCGTGGCCAGTTCCTCCGCGAGG 540  
 Db 498 GGAAGCAGAGAGAGCAAGCCCAAGCCGACAGCCGTGGCCGTGGCCAGTTCCTCCGCGAGG 557  
 Qy 541 GGCAGAGTGTCTGAGACTCGGGAGCCATTTGACCAATCGTCTCTGAGGATGGAGACTG 600  
 Db 558 GGCAGAGTGTCTGAGACTCGGGAGCCATTTGACCAATCGTCTCTGAGGATGGAGACTG 617  
 Qy 601 GTGAGCGGTGTCTGAGTGTCTCAGGAGAGTATTAACATCCCGAGCTCCAGTGGG 660  
 Db 618 GTGAGCGGTGTCTGAGTGTCTCAGGAGAGTATTAACATCCCGAGCTCCAGTGGG 677  
 Qy 661 CAAAGTCTCCATGGGTGCTGTATGAGGGCTGTGAGGGAGGAGAAAGCAGAGGAACTGCT 720  
 Db 678 CAAAGTCTCCATGGGTGCTGTATGAGGGCTGTGAGGGAGGAGAAAGCAGAGGAACTGCT 737  
 Qy 721 GTTGTACTCTGGAGCCCTCGAGGGGCTTCTCTCATCCGGAGAGCCAGAGAGG 780  
 Db 738 GTTGTACTCTGGAGCCCTCGAGGGGCTTCTCTCATCCGGAGAGCCAGAGAGG 797  
 Qy 781 CTCCTACTCTCTGCTCAGTCCGCTCAGCGGCTGTGATCTCTGGAGCCGATTCAGACACTA 840  
 Db 798 CTCCTACTCTCTGCTCAGTCCGCTCAGCGGCTGTGATCTCTGGAGCCGATTCAGACACTA 857  
 Qy 841 CAGGATCCACTGCTTGAACAATGGCTGTGATCATCTCACCGCGCTCACCTTCCCTC 900



Db 241 GGCCCCAAGCCCTAACCTGCTCCAGCCAGAGATCGCTCTCAGCAGAGCTGTCTTCCCA 300  
 QY 336 AGCTTTTGATGACAAACCAATTCCTCGATGATGTCTTCTGAGTGTCTGCTGAGGAA 395  
 Db 301 AGCTTTTGATGACAAACCAATTCCTCGATGATGTCTTCTGAGTGTCTGCTGAGGAA 360  
 QY 396 CAATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTG 455  
 Db 361 CAATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTG 420  
 QY 456 TCCAGGCCAGGACCTGTGACATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 515  
 Db 421 TCCAGGCCAGGACCTGTGACATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
 QY 516 TGGGCAAGTTTCCCGCAGTGTGCGCGGAGCTGTGCTGAGAGCTGCGGAGCCATGA 575  
 Db 481 TGGGCAAGTTTCCCGCAGTGTGCGCGGAGCTGTGCTGAGAGCTGCGGAGCCATGA 540  
 QY 576 CCATCGTCTCTGAGGATGAGAGTGTGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTG 635  
 Db 541 CCATCGTCTCTGAGGATGAGAGTGTGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTG 600  
 QY 636 ATAACTATCCCGCAGCTTCCAGCTGGGCAAGTCTCCATGGTGGCTGTATGAGGCTGA 695  
 Db 601 ATAACTATCCCGCAGCTTCCAGCTGGGCAAGTCTCCATGGTGGCTGTATGAGGCTGA 660  
 QY 696 GCAGGAGAGAAAGCAGAGGAATCTGTGTTTAACTGGGAAACCTGGAGGGGCTTCCCTCA 755  
 Db 661 GCAGGAGAGAAAGCAGAGGAATCTGTGTTTAACTGGGAAACCTGGAGGGGCTTCCCTCA 720  
 QY 756 TCCGGAGAGCCAGACAGAGAGCTCTTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 815  
 Db 721 TCCGGAGAGCCAGACAGAGAGCTCTTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
 QY 816 CATCTCTGGACCGGATCAGACACTACAGGATCCACTGCTTGAATGGCTGGCTGTACA 875  
 Db 781 CATCTCTGGACCGGATCAGACACTACAGGATCCACTGCTTGAATGGCTGGCTGTACA 840  
 QY 876 TCTCAGCGGCTCAGCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 935  
 Db 841 TCTCAGCGGCTCAGCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 900  
 QY 936 CGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCT 995  
 Db 901 CGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 996 CTGGCAGGATATACCCCTTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1055  
 Db 961 CTGGCAGGATATACCCCTTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 QY 1056 TGGACAGCTCCCTCTCTGTTTCTGAGGCTGCCAGGGGAGAGTCTCTTCTCAGTGAGG 1115  
 Db 1021 TGGACAGCTCCCTCTCTGTTTCTGAGGCTGCCAGGGGAGAGTCTCTTCTCAGTGAGG 1080  
 QY 1116 GTCTCGGAGTCCCTCAGCTTCTATCAGCTCAATGAGGAGCTCTCTCTTGGATG 1175  
 Db 1081 GTCTCGGAGTCCCTCAGCTTCTATCAGCTCAATGAGGAGCTCTCTCTTGGATG 1140  
 QY 1176 ATGCTTAG 1183  
 Db 1141 ATGCTTAG 1148

RESULT 5  
 AX511153  
 LOCUS AX511153  
 DEFINITION Sequence 4 from Patent WO0242452.  
 ACCESSION AX511153  
 VERSION AX511153.1 GI:23392046  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1  
 AUTHORS McElade, J.C. and Loreto, M.P.  
 TITLE Adapter gene  
 JOURNAL Patent: WO 0242452-A 4 30-MAY-2002;  
 The Hospital for Sick Children (CA)  
 FEATURES  
 source Location/Qualifiers  
 1..786  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 62.1%; Score 735; DB 6; Length 786;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 398 ATGGGAAGTCTGTCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAAGTTCCTCTGTC 457  
 Db 1 ATGGGAAGTCTGTCCAGCAGAGAAATCTCTGCAAGCCCAAGCTTGAAGTTCCTCTGTC 60  
 QY 458 GAGGCGAGGACCTGTGACCATGGAGCAGAGAGAAAGCAAGCCAGCGGTGGCCCTG 517  
 Db 61 CAAGGCCAGGACCTGTGACCATGGAGCAGAGAGAAAGCAAGCCAGCGGTGGCCCTG 120  
 QY 518 GGCAGTTCCTCCGCGCAGGTGGCCCGGAGCTGTGCTGAGACTCGGGGAGCAATGACC 577  
 Db 121 GGCAGTTCCTCCGCGCAGGTGGCCCGGAGCTGTGCTGAGACTCGGGGAGCAATGACC 180  
 QY 578 ATCGTCTCTGAGGATGGAGACTGTGGAGCTGTGCTGAGTCTCAGGAGGAGATAT 637  
 Db 181 ATCGTCTCTGAGGATGGAGACTGTGGAGCTGTGCTGAGTCTCAGGAGGAGATAT 240  
 QY 638 AACATCCCCAGGCTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 697  
 Db 241 AACATCCCCAGGCTCCAGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 300  
 QY 698 AGGGAAGAGCAGAGAACTCTGTGTACCTGGGAAACCTGGAGGGGCTTCTCTCATC 757  
 Db 301 AGGGAAGAGCAGAGAACTCTGTGTGTACCTGGGAAACCTGGAGGGGCTTCTCTCATC 360  
 QY 758 CGGAGAGCCAGACAGCAGAGAGGCTCTTACTCTCTGTGCTGCTGCTGCTGCTGCTGCA 817  
 Db 361 CGGAGAGCCAGACAGCAGAGAGGCTCTTACTCTCTGTGCTGCTGCTGCTGCTGCTGCA 420  
 QY 818 TCCTGGGACCGGATCAGACACTCAGGATCCACTGCTTGCATGCTGCTGCTGCTGCTGCT 877  
 Db 421 TCCTGGGACCGGATCAGACACTCAGGATCCACTGCTTGCATGCTGCTGCTGCTGCTGCT 480  
 QY 878 TCACCGGCTCAGCTTCCCTCTCACTCCAGGCTTGTGGACCAATTAATCTGAGCTGGCG 937  
 Db 481 TCACCGGCTCAGCTTCCCTCTCACTCCAGGCTTGTGGACCAATTAATCTGAGCTGGCG 540  
 QY 938 GATGACATCTGCTGCTCACTCAAGAGCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 997  
 Db 541 GATGACATCTGCTGCTCACTCAAGAGCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 998 GCAAGGATATACCCCTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057  
 Db 601 GCAAGGATATACCCCTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 QY 1058 GACAGCTCCCTCTCTGTTTCTGAGGCTGCCAGGGGAGGAGTCTCTCTCAGTGAGGGT 1117  
 Db 661 GACAGCTCCCTCTCTGTTTCTGAGGCTGCCAGGGGAGGAGTCTCTCTCAGTGAGGGT 720  
 QY 1118 CTCGGGAGTCCCTCAGCTTCTATCAGGCTTGAATGACGAGGCTGTCTCTTTGGATGAT 1177  
 Db 721 CTCGGGAGTCCCTCAGCTTCTATCAGGCTTGAATGACGAGGCTGTCTCTTTGGATGAT 780  
 QY 1178 GCCTAG 1183  
 Db 781 GCCTAG 786

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

RESULT 6
AX572845
LOCUS
DEFINITION
Sequence 1 from Patent WO20205707.
ACCESSION
AX572845
VERSION
AX572845.1 GI:26004935
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Holland, S.J., Mendenhall, M.K., Pardo, J., Spencer, C., Fu, A.C.,
Luo, Y., Payan, D.G., Mancebo, H.S., Wu, J., Zhou, X., Shen, M.,
Liao, X.C. and Sheng, N.
TITLE
Cloning of an inhibitor of antigen-receptor signaling by a
retroviral-based functional screen
JOURNAL
Patent: WO 02055707-A 1 18-JUL-2002;
Rigel Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source
1..786
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
1..786
/feature="unassigned protein product"
/codon_start=1
/protein_id="CAD58542.1"
/db_xref="GI:26004936"
/db_xref="REMBL:CAD58542"
/translation="MGSLLPSRRKSLPSLSVSSVQGGVPTWEAKRSKATAVALGSPF
AGPAELSLRGLGELPTIVSDGDMWTVLSVRSREYNIPSVHAKVSHGLYEGLSRE
KABELLLPQNGPAGFLIRSDTRGSYLSVRLSRPASWDRIRHYRTHCLDNGWLYI
SPRFTFSLQALVDHYSELADICCLKEPCVLORAPLPGKDIPLPTVQRTPLNWK
ELDSSLFSEAAATGEESLLSEGLSELSFYISLNDVAVSLDDA"
ORIGIN
Query Match 62.1%; Score 735; DB 6; Length 786;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGATCTCTGTC 457
DB 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGATCTCTGTC 60
QY 458 CAAGCCAGGACCTGTGACCATGCAAGCAGAGAGAGAGCAAGCCAGCCGCTGGCCCTG 517
DB 61 CAGGCCAGGACCTGTGACCATGCAAGCAGAGAGAGAGCAAGCCAGCCGCTGGCCCTG 120
QY 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCGTGAGACTCGGGGAGCCATTGACC 577
DB 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCGTGAGACTCGGGGAGCCATTGACC 180
QY 578 ATCGTCTCTGAGGATGGAGACTGTGTGACCGGTGTGTCTGAAGTCTCAGGCAGAGAGTAT 637
DB 181 ATCGTCTCTGAGGATGGAGACTGTGTGACCGGTGTGTCTGAAGTCTCAGGCAGAGTAT 240
QY 638 AACATCCCACGCTCCAGTGGGCAAGTCTCCATGGTGGTGTATGAGGCGCTGAGC 697
DB 241 AACATCCCACGCTCCAGTGGGCAAGTCTCCATGGTGGTGTATGAGGCGCTGAGC 300
QY 698 AGGAGAAAGCAGAGGAAGTCTCTGTTGTTACCTGGGAACCTCGAGGGGCGCTTCCTCATC 757
DB 301 AGGAGAAAGCAGAGGAAGTCTCTGTTGTTACCTGGGAACCTCGAGGGGCGCTTCCTCATC 360
QY 758 CGGGAGCCGACAGCAGAGAGGCTCTTACTCTGTGTCAGTCCGCTCAGCCGCCCTGCA 817
DB 361 CGGGAGCCGACAGCAGAGAGGCTCTTACTCTGTGTCAGTCCGCTCAGCCGCCCTGCA 420
QY 818 TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGACATGCGCTGAGCTGTGATCATC 877
DB 421 TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGACATGCGCTGAGCTGTGATCATC 480

```

```

QY 878 TCACCGGCGCTCACCTTCCCTCCTCCTCAGGCGCCCTGGTGGACCATTTACTCTGAGCTGGC 937
DB 481 TCACCGGCGCTCACCTTCCCTCCTCCTCAGGCGCCCTGGTGGACCATTTACTCTGAGCTGGC 540
QY 938 GATGACATCTGCTGCTTACTCTCAGAGGCCCTGTGCTCCCTGAGAGGCTGGCCCGCTCCCT 997
DB 541 GATGACATCTGCTGCTTACTCTCAGAGGCCCTGTGCTCCCTGAGAGGCTGGCCCGCTCCCT 600
QY 998 GGCAAGGATATACCCCTACCTGCTGAGCTGTGAGAGGACACCACTCAACTGGAAAGAGCTG 1057
DB 601 GGCAAGGATATACCCCTACCTGCTGAGCTGTGAGAGGACACCACTCAACTGGAAAGAGCTG 660
QY 1058 GAGAGTCCCTCTCTGTTTCTGAGCTGCCAGAGGAGGAGTCTCTCTCAGTCAAGGT 1117
DB 661 GAGAGTCCCTCTCTGTTTCTGAGCTGCCAGAGGAGGAGTCTCTCTCAGTCAAGGT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTTGAATGAGAGGCTGTCTCTTGGATCAT 1177
DB 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTTGAATGAGAGGCTGTCTCTTGGATCAT 780
QY 1178 GCCTAG 1183
DB 781 GCCTAG 786
RESULT 7
AF290985 786 bp mRNA linear PRI 21-JAN-2003
LOCUS
DEFINITION Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION
AF290985
VERSION
AF290985.1 GI:17351920
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 786)
Loreto, M.P. and McGlade, C.J.
Cloning and characterization of human Src-like adaptor protein 2
and a novel splice isoform, SLAP-2-v
Oncogene 22 (2), 266-273 (2003)
JOURNAL
MEDLINE
22415750
PUBMED
12527895
REFERENCE
2 (bases 1 to 786)
Loreto, M.P. and McGlade, C.J.
Direct Submission
Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
AUTHORS
JOURNAL
FEATURES
Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="20"
/tissue_type="thymus"
1..786
/feature="SLAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAL38197.1"
/db_xref="GI:17351921"
CDS
/translation="MGSLLPSRRKSLPSLSVSSVQGGVPTWEAKRSKATAVALGSPF
AGPAELSLRGLGELPTIVSDGDMWTVLSVRSREYNIPSVHAKVSHGLYEGLSRE
KABELLLPQNGPAGFLIRSDTRGSYLSVRLSRPASWDRIRHYRTHCLDNGWLYI
SPRFTFSLQALVDHYSELADICCLKEPCVLORAPLPGKDIPLPTVQRTPLNWK
ELDSSLFSEAAATGEESLLSEGLSELSFYISLNDVAVSLDDA"
ORIGIN
Query Match 62.1%; Score 735; DB 9; Length 786;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 398 ATGGGAAGTCTCCAGCAGAGAGAAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 457  
 Db 1 ATGGGAAGTCTCCAGCAGAGAGAAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 60  
 QY 458 CAAGGCCAGGACCTGTGACCATGAAGCAGAGAGAACCAAGGCCACAGCGTGGCCCTG 517  
 Db 61 CAAGGCCAGGACCTGTGACCATGAAGCAGAGAGAACCAAGGCCACAGCGTGGCCCTG 120  
 QY 518 GGCAGTTTCCCGGACAGGTGGCCCGCCGACAGGTGCTGCTGAGACTCGGGAGCCATTGACC 577  
 Db 121 GGCAGTTTCCCGGACAGGTGGCCCGCCGACAGGTGCTGCTGAGACTCGGGAGCCATTGACC 180  
 QY 578 ATCTGCTCTGAGATGAGAGACTGGTGGACGGTGTCTGCTGAAGCTCAGGCAGAGAGTAT 637  
 Db 181 ATCTGCTCTGAGATGAGAGACTGGTGGACGGTGTCTGCTGAAGCTCAGGCAGAGAGTAT 240  
 QY 638 AACATCCCAGCGCTCAGCTGGGCAAAATCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 697  
 Db 241 AACATCCCAGCGCTCAGCTGGGCAAAATCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 300  
 QY 698 AGGAGAAAGCAGAGAACTGCTGTGTTTACTCTGGAAACCTGGAGGGCTTCTCTCATC 757  
 Db 301 AGGAGAAAGCAGAGAACTGCTGTGTTTACTCTGGAAACCTGGAGGGCTTCTCTCATC 360  
 QY 758 CCGGAGAGCCAGACAGGAGAGGCTTCTTCTCTGTCAGTCCGCTCAGCCGCTGCA 817  
 Db 361 CCGGAGAGCCAGACAGGAGAGGCTTCTTCTCTGTCAGTCCGCTCAGCCGCTGCA 420  
 QY 818 TCTTGGACCCGATCAGACACTACAGGATCCACTGCTTGAACATGCTGCTGTACATC 877  
 Db 421 TCTTGGACCCGATCAGACACTACAGGATCCACTGCTTGAACATGCTGCTGTACATC 480  
 QY 878 TCACGCGGCTCAGCTTCCCTCAGTCCAGGCGCTGTCTGCTGAGAGGCTGCGCCCTCCCT 937  
 Db 481 TCACGCGGCTCAGCTTCCCTCAGTCCAGGCGCTGTCTGCTGAGAGGCTGCGCCCTCCCT 540  
 QY 938 GATGACATCTGCTGCTACTCAAGAGGCGCTGTCTGCTGAGAGGCTGCGCCCTCCCT 997  
 Db 541 GATGACATCTGCTGCTACTCAAGAGGCGCTGTCTGCTGAGAGGCTGCGCCCTCCCT 600  
 QY 998 GGCAGGATATACCCCTACCTGTGCTGCTGAGAGGACCACTCACTGAAAGAGCTG 1057  
 Db 601 GGCAGGATATACCCCTACCTGTGCTGCTGAGAGGACCACTCACTGAAAGAGCTG 660  
 QY 1058 GACAGCTCCCTCTCTTCTGAGCTCCACAGAGGAGGCTCTCTCTCAGTGAAGGT 1117  
 Db 661 GACAGCTCCCTCTCTTCTGAGCTCCACAGAGGAGGCTCTCTCTCAGTGAAGGT 720  
 QY 1118 CTCGCGGAGTCCCTCAGCTTCTATCATGAGCTGATGAGAGGCTGTCTCTTGGATGAT 1177  
 Db 721 CTCGCGGAGTCCCTCAGCTTCTATCATGAGCTGATGAGAGGCTGTCTCTTGGATGAT 780  
 QY 1178 GCCTAG 1183  
 Db 781 GCCTAG 786

RESULT 8  
 AF326353  
 LOCUS 786 bp mRNA linear PRI 08-NOV-2001  
 DEFINITION Homo sapiens Src-like adapter protein-2 mRNA, complete cds.  
 ACCESSION AF326353  
 VERSION AF326353.1 GI:16797891  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 786)  
 Holland, S.J., Liao, X.C., Mendenhall, M.K., Zhou, X., Pardo, J.,  
 Chu, P., Spencer, C., Fu, A.C., Sheng, N., Yu, P., Pali, E., Nagin, A.,  
 Shen, M., Yu, S., Chan, E., Wu, X., Li, C., Weisbach, M.,  
 Aversa, G., Kolbinger, F., Bennett, M.K., Molineaux, S., Luo, Y.,

TITLE Pavan, D.G., Mancebo, H.S.Y. and Wu, J.  
 JOURNAL Functional Cloning of Src-like Adapter Protein-2 (SLAP-2), a Novel  
 MEDLINE Inhibitor of Antigen Receptor Signaling  
 PUBMED J. Exp. Med. 194 (9), 1263-1276 (2001)  
 REFERENCE 2153359  
 2 (bases 1 to 786)  
 TITLE Direct Submission  
 JOURNAL Submitted (05-DEC-2000) Rigol Pharmaceutical Inc., 240 East Grand  
 AVENUE, South San Francisco, CA 94080, USA  
 FEATURES Location/Qualifiers  
 source 1..786  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 CDS 1..786  
 /note="SLAP2; SH2 and SH3 domain-containing adaptor"  
 /codon\_start=1  
 /product="Src-like adapter protein-2"  
 /protein\_id="AAU29204.1"  
 /db\_xref="GI:16797892"  
 /translation="MGSLPERRKSLPSLSVQGGPVTMEASRKATAVALGSPF  
 AGGPAELSLRLGPELTVSDGDMWTVLSEVSGREYNIPSHVAKYHGMVGLSRE  
 KARELLLPNGSGAPLIRESDTRGYSLSVLSPEASWDRIRHRIHCLDNGWLYI  
 SPRLTFPSLQALVDHYSELADDDICLLKSPVLRAGPLPGKIDPLPVTVQRTPLNWK  
 ELDSLLPSEANYGERSLLSEGLRESLSFYISLNDSEAVSLDDA"

## ORIGIN

Query Match 62.1%; Score 735; DB 9; Length 786;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 398 ATGGGAAGTCTGCCAGCAGAGAAATCTCTCCAGCCCAAGCTTGAGTTCTCTGTC 457  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTCCAGCCCAAGCTTGAGTTCTCTGTC 60  
 QY 458 CAAGGCCAGGACCTGTGACCATGAAGCAGAGAGAACCAAGGCCACAGCGTGGCCCTG 517  
 Db 61 CAAGGCCAGGACCTGTGACCATGAAGCAGAGAGAACCAAGGCCACAGCGTGGCCCTG 120  
 QY 518 GGCAGTTTCCCGGACAGGTGGCCCGCCGACAGGTGCTGCTGAGACTCGGGAGCCATTGACC 577  
 Db 121 GGCAGTTTCCCGGACAGGTGGCCCGCCGACAGGTGCTGCTGAGACTCGGGAGCCATTGACC 180  
 QY 578 ATCTGCTCTGAGATGAGAGACTGGTGGACGGTGTCTGCTGAAGCTCAGGCAGAGAGTAT 637  
 Db 181 ATCTGCTCTGAGATGAGAGACTGGTGGACGGTGTCTGCTGAAGCTCAGGCAGAGAGTAT 240  
 QY 638 AACATCCCAGCGTCCAGCTGGGCAAAATCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 697  
 Db 241 AACATCCCAGCGTCCAGCTGGGCAAAATCTCCCATGGGTGGCTGTATGAGGGCCTGAGC 300  
 QY 698 AGGAGAAAGCAGAGAACTGCTGTGTTTACTCTGGAAACCTGGAGGGCTTCTCTCATC 757  
 Db 301 AGGAGAAAGCAGAGAACTGCTGTGTTTACTCTGGAAACCTGGAGGGCTTCTCTCATC 360  
 QY 758 CCGGAGAGCCAGACAGGAGAGGCTTCTTCTCTGTCAGTCCGCTCAGCCGCTGCA 817  
 Db 361 CCGGAGAGCCAGACAGGAGAGGCTTCTTCTCTGTCAGTCCGCTCAGCCGCTGCA 420  
 QY 818 TCCTGGGACCCGATCAGACACTACAGGATCCACTGCTTGAACATGCTGCTGTACATC 877  
 Db 421 TCCTGGGACCCGATCAGACACTACAGGATCCACTGCTTGAACATGCTGCTGTACATC 480  
 QY 878 TCACGCGGCTCAGCTTCCCTCAGTCCAGGCGCTGTCTGCTGAGAGGCTGCGCCCTCCCT 937  
 Db 481 TCACGCGGCTCAGCTTCCCTCAGTCCAGGCGCTGTCTGCTGAGAGGCTGCGCCCTCCCT 540  
 QY 938 GATGACATCTGCTGCTACTCAAGAGGCGCTGTCTGCTGAGAGGCTGCGCCCTCCCT 997

```

Db      541 GATGACATCTGCTGCTACTCAAGAGCCCTGTCTCTGTCAGAGGGCTGCGCCGCTCCCT 600
QY      998 GCGAGGATATACCCCTACCTGTGACTGTGTCAGAGGACACCACTCACTGGAAGAGCTG 1057
Db      601 GCGAGGATATACCCCTACCTGTGACTGTGTCAGAGGACACCACTCACTGGAAGAGCTG 660
QY      1058 GACAGCTCCCTCTGTTTCTGAAGTGCACACAGGGAGGAGTCTCTTCTCAGTGAGGGT 1117
Db      661 GACAGCTCCCTCTGTTTCTGAAGTGCACACAGGGAGGAGTCTCTTCTCAGTGAGGGT 720
QY      1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATCAGAGGCTGCTCTTTGGATGAT 1177
Db      721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATCAGAGGCTGCTCTTTGGATGAT 780
QY      1178 GCCTAG 1183
Db      781 GCCTAG 786

RESULT 9
AK025645
LOCUS      2415 bp mRNA linear PRI 13-SEP-2003
DEFINITION Homo sapiens cDNA: FL221992 fis, clone HEP06554.
ACCESSION AK025645
VERSION    AK025645.1 GI:10438227
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
           Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
           Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
           NEDO human cDNA sequencing project
           Unpublished
           2 (bases 1 to 2415)
           Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
           Shibahara,T., Tanaka,T. and Nakamura,Y.
           Direct Submission
           Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
           University of Tokyo, Laboratory of Genome Structure Analysis, Human
           Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
           Japan (E-mail: ficdn@aims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
           Fax:81-3-5449-5416)
           NEDO human cDNA sequencing project supported by Ministry of
           International Trade and Industry of Japan: cDNA full insert
           sequencing: Research Association for Biotechnology; cDNA library
           construction, 5'- & 3'- and one pass sequencing: Department of
           Virology and Human Genome Center, Institute of Medical Science,
           University of Tokyo (partly supported by Science and Technology
           Agency).
           Location/Qualifiers
             1..2415
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="HEP06554"
               /cell_line="HepG2"
               /clone_lib="HEP"
               /note="cloning vector pME185FL3"
             61..846
               /note="unnamed protein product"
               /codon_start=1
               /protein_id="BAB15201.1"
               /db_xref="GI:10438228"
               /translation="MGSPLSRKSLPSLSVSGQGPVMEERSKATAVALGSP
               AGGPAELSLRGEPLTVSEGDWTVLSEVSGREYNIPSVHAKVSHGLVYGLSRE
               KAEELLPGNPGAFILRESQTRGYSLSVRLSRFASMDRIHRYHICLDNGWLYI
               SPRLTFPQLAIVHYSLEADICLLKEPCVLRAGPLPGKDIPLPTVQRTPLNWK
               ELDSLLFPSEATGEESLLSGLRESUSFYISLNDVAVSLDDA"

FEATURES
         source
           1..2415
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="HEP06554"
             /cell_line="HepG2"
             /clone_lib="HEP"
             /note="cloning vector pME185FL3"
           61..846
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="BAB15201.1"
             /db_xref="GI:10438228"
             /translation="MGSPLSRKSLPSLSVSGQGPVMEERSKATAVALGSP
             AGGPAELSLRGEPLTVSEGDWTVLSEVSGREYNIPSVHAKVSHGLVYGLSRE
             KAEELLPGNPGAFILRESQTRGYSLSVRLSRFASMDRIHRYHICLDNGWLYI
             SPRLTFPQLAIVHYSLEADICLLKEPCVLRAGPLPGKDIPLPTVQRTPLNWK
             ELDSLLFPSEATGEESLLSGLRESUSFYISLNDVAVSLDDA"
         CDS
           1..2415
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="HEP06554"
             /cell_line="HepG2"
             /clone_lib="HEP"
             /note="cloning vector pME185FL3"
           61..846
             /note="unnamed protein product"
             /codon_start=1
             /protein_id="BAB15201.1"
             /db_xref="GI:10438228"
             /translation="MGSPLSRKSLPSLSVSGQGPVMEERSKATAVALGSP
             AGGPAELSLRGEPLTVSEGDWTVLSEVSGREYNIPSVHAKVSHGLVYGLSRE
             KAEELLPGNPGAFILRESQTRGYSLSVRLSRFASMDRIHRYHICLDNGWLYI
             SPRLTFPQLAIVHYSLEADICLLKEPCVLRAGPLPGKDIPLPTVQRTPLNWK
             ELDSLLFPSEATGEESLLSGLRESUSFYISLNDVAVSLDDA"

```

## ORIGIN

```

Query Match      62.1%; Score 735; DB 9; Length 2415;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      398 ATGGGAAGTCTGCCCCAGCAGAGAAATCTCTGCCAGCCCCCAAGCTTCAGTTCTCTCTGTC 457
Db      61 ATGGGAAGTCTGCCCCAGCAGAGAAATCTCTGCCAGCCCCCAAGCTTCAGTTCTCTCTGTC 120
QY      458 CAAAGCCAGGAGCCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGCCCCCTG 517
Db      121 CAAAGCCAGGAGCCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCCGTGCCCCCTG 180
QY      518 GGCAGTTTCCCGGAGGTGGCCGCGCGAGCTGTGCTGAGACTCGGGAGGCCATTGACC 577
Db      181 GGCAGTTTCCCGGAGGTGGCCGCGCGAGCTGTGCTGAGACTCGGGAGGCCATTGACC 240
QY      578 ATCGTCTCTGAGGATGAGACTGTGTGACGCGTGTCTGAAAGTCTCAGGCAGAGAGTAT 637
Db      241 ATCGTCTCTGAGGATGAGACTGTGTGACGCGTGTCTGAAAGTCTCAGGCAGAGAGTAT 300
QY      638 AACATCCCCAGCGTCCAGCTGGGCAAAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 697
Db      301 AACATCCCCAGCGTCCAGCTGGGCAAAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 360
QY      698 AGGAGAAAGCAGAGGAACTGCTGTTGTTACCTGGGAAACCTTGGAGGGGCTTCTCTCATC 757
Db      361 AGGAGAAAGCAGAGGAACTGCTGTTGTTACCTGGGAAACCTTGGAGGGGCTTCTCTCATC 420
QY      758 CGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCCCTGCA 817
Db      421 CGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCCCTGCA 480
QY      818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCACATGGCTGGCTGTACATC 877
Db      481 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCACATGGCTGGCTGTACATC 540
QY      878 TCACCGGCTCTACCTTCCCTCCTCCTCAGGAGGCTGTGCTCTGTCAGAGGGCTGCCCCCTCCCT 937
Db      541 TCACCGGCTCTACCTTCCCTCCTCCTCAGGAGGCTGTGCTCTGTCAGAGGGCTGCCCCCTCCCT 600
QY      938 GATGACATCTGCTGCTTACTCAAGGAGCCCTGTGCTCTGTCAGAGGGCTGCCCCCTCCCT 997
Db      601 GATGACATCTGCTGCTTACTCAAGGAGCCCTGTGCTCTGTCAGAGGGCTGCCCCCTCCCT 660
QY      998 GGCAGAGATATACCTTACTGTCAGAGGACACCACTCAACTGGAAAGAGCTG 1057
Db      661 GGCAGAGATATACCTTACTGTCAGAGGACACCACTCAACTGGAAAGAGCTG 720
QY      1058 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGT 1117
Db      721 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGT 780
QY      1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGAGAGGCTGTCTCTTTGGATGAT 1177
Db      781 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGAGAGGCTGTCTCTTTGGATGAT 840
QY      1178 GCCTAG 1183
Db      841 GCCTAG 846

RESULT 10
AX780857
LOCUS      2788 bp DNA linear PAT 14-JUL-2003
DEFINITION Sequence 3014 from Patent WO03039443.
ACCESSION AX780857
VERSION    AX780857.1 GI:32697851
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE 1  
 AUTHORS Haerlrich, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S., Dugas, M., Eils, R., Brors, B. and Mergenthaler, S.  
 TITLE Novel genetic markers for leukemias  
 JOURNAL Patent: WO 03039443-A 3014 15-MAY-2003;  
 Deutsches Krebsforschungszentrum (DE);  
 Ludwig-Maximilians-Universität München (DE);  
 PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)  
 FEATURES  
 source 1. .2788  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 57.8%; Score 684; DB 6; Length 2788;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 784; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 398 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 457  
 Db 387 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 446  
 QY 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 517  
 Db 447 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 506  
 QY 518 GGCAGTTTCGCCGAGTGGCCCGGCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 577  
 Db 507 GGCAGTTTCGCCGAGTGGCCCGGCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 566  
 QY 578 ATGCTCTGAGATGGAGCTGTGGACGCTGTGTGTAAGTCTCAGGAGAGAGTAT 637  
 Db 567 ATGCTCTGAGATGGAGCTGTGGACGCTGTGTGTAAGTCTCAGGAGAGAGTAT 626  
 QY 638 AACATCCCAAGCTCCAGTGGGCAAGTCTCCCATGGGTGCTGTATGAGGGCTGAGC 697  
 Db 627 AACATCCCAAGCTCCAGTGGGCAAGTCTCCCATGGGTGCTGTATGAGGGCTGAGC 686  
 QY 698 AGGAGAAAGCAGAGAACTGTGTTACTTGGAAACCTTGGAGGGCCCTTCTCATC 757  
 Db 687 AGGAGAAAGCAGAGAACTGTGTTACTTGGAAACCTTGGAGGGCCCTTCTCATC 746  
 QY 758 CGGAGAGCCAGACACAGGAGAGGCTCTTACTCTCTGTGAGTCCGCTCAGCGCCCTGCA 817  
 Db 747 CGGAGAGCCAGACACAGGAGAGGCTCTTACTCTCTGTGAGTCCGCTCAGCGCCCTGCA 806  
 QY 818 TCCTGGACCGGATCAGACACTACAGGATCCACTGCTTGACATGGCTGGCTGTACATC 877  
 Db 807 TCCTGGACCGGATCAGACACTACAGGATCCACTGCTTGACATGGCTGGCTGTACATC 866  
 QY 878 TCACCGCGCTCACCTTCCCTCCTCAGGCGCCCTGGTGACCACTTACTCTGAGTGGCG 937  
 Db 867 TCACCGCGCTCACCTTCCCTCCTCAGGCGCCCTGGTGACCACTTACTCTGAGTGGCG 926  
 QY 938 GATGACATCTGCTCTACTCAGAGGACCTGTGTCTCTGAGAGGCTGGCCCGCTCCCT 997  
 Db 927 GATGACATCTGCTCTACTCAGAGGACCTGTGTCTCTGAGAGGCTGGCCCGCTCCCT 986  
 QY 998 GGCAGGATATACCCCTACCTGTGACTGTGCGAGAGACACCACTCAACTGGAAGAGCTG 1057  
 Db 987 GGCAGGATATACCCCTACCTGTGACTGTGCGAGAGACACCACTCAACTGGAAGAGCTG 1046  
 QY 1058 GACAGTCCCTCTGTTCTTGAAGTGCACAGGGAGGAGTCTTCTCAGTGGGGT 1117  
 Db 1047 GACAGTCCCTCTGTTCTTGAAGTGCACAGGGAGGAGTCTTCTCAGTGGGGT 1106  
 QY 1118 CTCGGAGTCCCTCAGCTTCTACATCAGCCTCAATGAGGAGGCTGTCTCTTTGGATGAT 1177  
 Db 1107 CTCGGAGTCCCTCAGCTTCTACATCAGCCTCAATGAGGAGGCTGTCTCTTTGGATGAT 1166  
 QY 1178 GCCTAG 1183

Db 1167 GCCTAG 1172  
 RESULT 11  
 LOCUS AX511155  
 DEFINITION Sequence 6 from Patent WO0242452.  
 ACCESSION AX511155  
 VERSION AX511155.1 GI:23392047  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Mcglade, J.C. and Loreto, M.P.  
 TITLE Adapter gene  
 JOURNAL Patent: WO 0242452-A 6 30-MAY-2002;  
 The Hospital for Sick Children (CA)  
 FEATURES  
 source 1. .737  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 40.8%; Score 483; DB 6; Length 737;  
 Best Local Similarity 99.8%; Pred. No. 1.9e-268;  
 Matches 533; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 398 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 457  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60  
 QY 458 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 517  
 Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 120  
 QY 518 GGCAGTTTCGCCGAGTGGCCCGGCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 577  
 Db 121 GGCAGTTTCGCCGAGTGGCCCGGCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180  
 QY 578 ATGCTCTGAGATGGAGCTGTGTTACTTGGAAACCTTGGAGGGCCCTTCTCATC 637  
 Db 181 ATGCTCTGAGATGGAGCTGTGTTACTTGGAAACCTTGGAGGGCCCTTCTCATC 240  
 QY 638 AACATCCCAAGCTCCAGTGGGCAAGTCTCCCATGGGTGCTGTATGAGGGCTGAGC 697  
 Db 241 AACATCCCAAGCTCCAGTGGGCAAGTCTCCCATGGGTGCTGTATGAGGGCCCTGAGC 300  
 QY 698 AGGAGAAAGCAGAGAACTGTGTTACTTGGAAACCTTGGAGGGCCCTTCTCATC 757  
 Db 301 AGGAGAAAGCAGAGAACTGTGTTACTTGGAAACCTTGGAGGGCCCTTCTCATC 360  
 QY 758 CGGAGAGCCAGACACAGGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCGCCCTGCA 817  
 Db 361 CGGAGAGCCAGACACAGGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCGCCCTGCA 420  
 QY 818 TCCTGGACCGGATCAGACACTACAGGATCCACTGCTTGACATGGCTGGCTGTACATC 877  
 Db 421 TCCTGGACCGGATCAGACACTACAGGATCCACTGCTTGACATGGCTGGCTGTACATC 480  
 QY 878 TCACCGCGCTCACCTTCCCTCCTCAGGCGCCCTGGTGACCACTTACTCTGAG 931  
 Db 481 TCACCGCGCTCACCTTCCCTCCTCAGGCGCCCTGGTGACCACTTACTCTGAG 934  
 RESULT 12  
 AF290986  
 LOCUS  
 DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,  
 complete cds; alternatively spliced.  
 AF290986 737 bp mRNA linear PRI 21-JAN-2003  
 Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,  
 complete cds; alternatively spliced.



[illegible]

Db 11880 CTTTCCTCCTCGCTGGCTGTGCTTTGGAGGGTTCCCCAGTCCAGATCCCTAAGGAG 11821

Qy 121 CATGGGCGAGCTGATCCATCCCTGGTGTACAACACTGCTGACTGCAGACAGATGCTGAGCT 180

Db 11820 CATGGGCGAGCTGATCCATCCCTGGTGTACAACACTGCTGACTGCAGACAGATGCTGAGCT 11761

Qy 181 ACCCAACCAACACCTAGCTCTCCCTGAGATCTCCCTGAGATCTCCAGGCTGAGAGGTTCTGGGTG 240

Db 11760 ACCCAACCAACACCTAGCTCTCCCTGAGATCTCCCTGAGATCTCCAGGCTGAGAGGTTCTGGGTG 11701

Qy 241 TCCTAGGACCAAGGACACTGGCAGACTTCACAGAGGGGCCCCAAAGCCCTAACCTGTCCA 300

Db 11700 TCCTAGGACCAAGGACACTGGCAGACTTCACAGAGGGGCCCCAAAGCCCTAACCTGTCCA 11641

Qy 301 GCCAGAGCATGCGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCA 354

Db 11640 GCCAGAGCATGCGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCA 11587

RESULT 14

AC026539

LOCUS

DEFINITION Homo sapiens chromosome 20 clone RP11-712N14 map 20, WORKING DRAFT

AC026539

VERSION

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE

ORGANISM Homo sapiens

REFERENCE

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F., Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G., Campatiano,A., Castle,A., Chospel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliveira,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanli,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Apr 27, 2000 this sequence version replaced gi:7283243.

All repeats were identified using RepeatMasker:

Smit,A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7115

Center clone name: 712\_N14

----- Summary Statistics

```

* 76825 81179: contig of 4355 bp in length
* 81180 81279: gap of 100 bp
* 81280 87009: contig of 5730 bp in length
* 87010 87109: gap of 100 bp
* 87110 90855: contig of 3746 bp in length
* 90856 90955: gap of 100 bp
* 90956 96200: contig of 5565 bp in length
* 96201 96200: gap of 100 bp
* 96201 102321: contig of 5701 bp in length
* 102322 102421: gap of 100 bp
* 102422 108293: contig of 5872 bp in length
* 108294 108393: gap of 100 bp
* 108394 116689: contig of 8296 bp in length
* 116690 116789: gap of 100 bp
* 116790 125364: contig of 8475 bp in length
* 125365 125364: gap of 100 bp
* 125365 136354: contig of 10990 bp in length
* 136355 136455: gap of 100 bp
* 136455 145833: contig of 9379 bp in length.
FEATURES
    source
        1..145833
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /map="20"
            /clone="RP11-712N14"
            /clone_lib="RPC1-11 Human Male BAC"
            1..1214
                /note="assembly_fragment"
            1315..2673
                /note="assembly_fragment"
            2774..4520
                /note="assembly_fragment"
            4621..5961
                /note="assembly_fragment"
            6062..7719
                /note="assembly_fragment"
            7820..9799
                /note="assembly_fragment"
            9900..11434
                /note="assembly_fragment"
            11535..14382
                /note="assembly_fragment"
            14483..17148
                /note="assembly_fragment"
            17249..19239
                /note="assembly_fragment"
            19340..21102
                /note="assembly_fragment"
            21203..23371
                /note="assembly_fragment"
            23472..25782
                /note="assembly_fragment"
            25883..28824
                /note="assembly_fragment"
            28925..31619
                /note="assembly_fragment"
            31720..34679
                /note="assembly_fragment"
            34780..37527
                /note="assembly_fragment"
            37628..40146
                /note="assembly_fragment"
            40247..43743
                /note="assembly_fragment"
            43844..47133
                /note="assembly_fragment"
            47234..51023
                /note="assembly_fragment"
            51124..54935
                /note="assembly_fragment
            clone_end:17

```

```

misc_feature      vector_side:right
55036..59553
Query Match      29.9%; Score 354; DB 2; Length 145833;
Best Local Similarity 100.0%; Pred. NO. 1.1e-193;
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGCTAGAGCTCCAAAGGACCCACACGCTGTGTCTGTGTGACAGAGCTCAAGGCGCCCTGGG 60
DB 99745 AGCTAGAGCTCCAAAGGACCCACACGCTGTGTCTGTGTGACAGAGCTCAAGGCGCCCTGGG 99804
QY 61 CTTTCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 120
DB 99805 CTTTCCTCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 99864
QY 121 CATGGGGAGCTGATCCATCCCTGGCTGTGTAACAATCTGTGCTGCTGCTGCTGCTGCTGCT 180
DB 99865 CATGGGGAGCTGATCCATCCCTGGCTGTGTAACAATCTGTGCTGCTGCTGCTGCTGCT 99924
QY 181 ACCCAAAACCAACACTAGCTCTCCCTGGAAGATCTCTCCAGGCTTGAGAGAGTCTTGGTG 240
DB 99925 ACCCAAAACCAACACTAGCTCTCCCTGGAAGATCTCTCCAGGCTTGAGAGAGTCTTGGTG 99984
QY 241 TCCTAGAGACCAAGGACACTGGCAGACTTCAGAGAGGCGCCCAAGGCTTAACCTGTCCA 300
DB 99985 TCCTAGAGACCAAGGACACTGGCAGACTTCAGAGAGGCGCCCAAGGCTTAACCTGTCCA 100044
QY 301 GCCAGAGCATCGCTCTCAGCAGAGCTCTCTTCCCAAGCCTTTGATGACAAACCA 354
DB 100045 GCCAGAGCATCGCTCTCAGCAGAGCTCTCTTCCCAAGCCTTTGATGACAAACCA 100098
RESULT 15
HSDJ977B1/c
LOCUS
DEFINITION
    Human DNA sequence from clone RPs-977B1 on chromosome 20. Contains
    ESTs, STSs, GSSs and three putative CpG islands. Contains the 3'
    end of the gene for a novel protein tyrosine kinase, a gene for
    heterogeneous nuclear ribonucleoprotein A3 pseudogene, the gene for
    RAB5-interacting protein, the TGF2 gene for TGF(beta)-induced
    transcription factor 2 with two isoforms, the MYRL2 gene for myosin
    regulatory light chain 2 (smooth muscle isoform), the 3' end of the
    gene KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4)
    with two isoforms and a novel gene, complete sequence.
ACCESSION
    AL050318
VERSION
    AL050318.13 GI:9581785
KEYWORDS
    HTG; CpG island; heterogeneous ribonucleoprotein; KIAA0964; myosin
    domain; TGF2; transcription factor; tyrosine kinase.
SOURCE
    Homo sapiens (human)
ORGANISM
    Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 145068)
Lloyd.D.
Direct Submission
Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 28, 2000 this sequence version replaced gi:5924017.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human

```

chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence is the entire insert of clone RP5-977B1. The true left end of clone RP2-460J8 is at 62245 in this sequence. The true right end of clone CDB-218229 is at 62218 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-977B1 is from the library RPI-3 constructed by the group of Pister de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: PCYPAC2.  
Location/Qualifiers

1. 145068  
/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="20"

/clone="RP5-977B1"

/clone\_lib="RPI-5"

568..875

/note="Alusx repeat: matches 1..311 of consensus"

1024..1580

/note="match: GSS: Em:AQ620498"

1047..1629

/note="match: GSS: Em:AQ779162"

1631..1923

/note="Alusx repeat: matches 1..294 of consensus"

2776..2842

/note="L2 repeat: matches 2357..2423 of consensus"

3891..3949

/note="MIR repeat: matches 34..90 of consensus"

4458..4611

/note="MIR repeat: matches 67..231 of consensus"

5133..5549

/note="L1MD2 repeat: matches 5924..6340 of consensus"

5550..5680

/note="FLAM\_C repeat: matches 1..132 of consensus"

5703..5890

/note="Alusq repeat: matches 120..308 of consensus"

5923..6012

/note="L1MC/D repeat: matches 5432..5528 of consensus"

6588..6889

/note="Alusp repeat: matches 1..303 of consensus"

join(7423..7784,9960..10039,10306..10394,10917..11329,34968..35059,36569..36724,37531..39351)

/gene="KIAA0964"

/product="dJ977B1.6.2 (KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4))"

/note="match: cDNAs: Em:AB023181"

/evidence="not experimental"

join(<7423..7784,10306..10394,10917..11329,34968..35059,36569..36724,37531..39351)

/gene="KIAA0964"

/product="dJ977B1.6.1 (KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4))"

/note="match: cDNAs: Em:U67140 Em:AB000277 Em:AB000276 Em:AB000275 Em:AB023181 Em:AF009204 Em:AB003594 Em:U67988 Em:U67137 Em:U67138 Em:U67987 Em:U67139 Em:AF131778 match: ESTs: Em:AI811867 Em:AI620816 Em:AA496750 Em:AA647532 Em:AI747584 Em:AA931146 Em:AI687761 Em:AA154844 Em:AA278057 Em:AI2020736 Em:AI497978 Em:AA191427 Em:AA611645 Em:AA760107 Em:AA959160 Em:AA466940 Em:AI80916 Em:U67912 Em:AA582202 Em:AI748854 Em:AI162233 Em:AI840413 Em:AI347137 Em:AA571014

FEATURES  
source

repeat\_region

misc\_feature

misc\_feature

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

Em:AA870776 Em:AA144662 Em:AA126125 Em:AA496805 Em:AA62099 Em:AA443364 Em:AA874407 Em:AI986170 Em:AI840039 Em:AA943725 Em:AA95674 Em:AI93631 Em:AA760121 Em:AA050197 Em:R4650 Em:R35758 Em:R42663 Em:AA433898 Em:R4251 Em:AI918247 Em:R19230 Em:R48469 Em:AA861813 Em:AA602005 Em:AA824392 Em:R13930 Em:AI948469 Em:AA861813 Em:AA602005 Em:AI201639 Em:W45187 Em:Z45916 Em:R12931 Em:AI936733 Em:R12936 Em:N36776 Em:R12916 Em:R13255 Em:AI983945 Em:AI126063 Em:T57999 Em:AI916248 Em:R75144 Em:W07614 Em:AI972603 Em:AA434575 Em:AI500286 Em:AI201649 Em:AI143454 Em:AI128723 Em:AI179528 Em:AA031035 Em:T08824 Em:N22262 Em:R51856 Em:R96084 Em:Z42146 Em:T49317 Em:R6101 Em:R41361 Em:AA064703 Em:AA295898 Em:AL049119 Em:AA222576 Em:R28675 Em:AI652021 Em:AA082413 Em:H05501 Em:N44297 Em:AA046329 Em:N45937 Em:AA577212 Em:AA610855 Em:AI017860 Em:W58510 Em:T08842 Em:AI057148 Em:N22846 /evidence="not experimental" join(<7423..7784,9960..10039,10917..11329,34968..35059,36569..36724,37531..37749) /gene="KIAA0964" /note="isoform 2" continues in dJ996C2 (AL445705) and BK2182L9 (AL390374) match: proteins: Tr:Q9Y2H0" /codon\_start=3 /evidence="not experimental" /product="dJ977B1.6.2 (KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4))" /protein\_id="CAC16960.1" /db\_xref="GI:11323333" /db\_xref="GOA:Q9Y2H0" /db\_xref="SWISS-PROT:Q9Y2H0" /translation="SSCLVAYKTPPPVPPRTTSTKPFISVTVQSTSTSAQTYLDSQD HKSEVTSQSLNSDSDLSSTRPSPVTRGVGVAPEAPPPPPKAAKSKQGLTSS ESHPEAPKRLSSIGIERTRRNGSHLSIDNGPKAIDVMAPESSDVSMSRRD TSDQDANDSSCKSERSLPDCPTPHNSISIDAGPROAKIAQIKNLSYGDNSDDA LEASLPDPWLETSSSPAPQAPQACRDGWFLLQALQETERLEGWCCMDKET KENNLSEELVGLKVSAGVSAOLLSQKQFQGLCEONLNDANPRPTAQLAGFMDI LQSLIEDISMKFDLHLKANSWQLVETPEKREKEKPPPPVPPKPAKSAFVSERDKA SASDKQQRKRLAAKRAASVRQNSATSAESIYVPEAQL" join(<7423..7784,10306..10394,10917..11329,34968..35059,36569..36724,37531..37749) /gene="KIAA0964" /note="isoform 1" continues in dJ996C2 (AL445705) and BK2182L9 (AL390374) match: proteins: Tr:O14488 Tr:O14564 Tr:O14489 Tr:P97839 Tr:O14490 Tr:P97836 Tr:P97837 Tr:O54773 Tr:P97841 Tr:P78335 /codon\_start=3 /evidence="not experimental" /product="dJ977B1.6.1 (KIAA0964 (ortholog of rat PSD-95/SAP90-associated protein 4))" /protein\_id="CAB75370.1" /db\_xref="GI:6983730" /db\_xref="GOA:Q9Y2H0" /db\_xref="SWISS-PROT:Q9Y2H0" /translation="SSCLVAYKTPPPVPPRTTSTKPFISVTVQSTSTSAQTYLDSQD HKSEVTSQSLNSDSDLSSTRPSPVTRGVGVAPEAPPPPPKAAKSKQGLTSS ESHPEAPKRLSSIGIERTRRNGSHLSIDNGPKAIDVMAPESSDVSMSRRD TSDQDANDSSCKSERSLPDCPTPHNSISIDAGPROAKIAQIKNLSYGDNS DDALPASLPDPWLETSSSPAPQAPQACRDGWFLLQALQETERLEGWCCMD KETENNLSEELVGLKVSAGVSAOLLSQKQFQGLCEONLNDANPRPTAQLAGFMDI LQSLIEDISMKFDLHLKANSWQLVETPEKREKEKPPPPVPPKPAKSAFVSERDKA SASDKQQRKRLAAKRAASVRQNSATSAESIYVPEAQL" 1065..10112 /note="6 copies 8 mer cgcctgc 79? conserved" 11537..11753 /note="MIR repeat: matches 11..236 of consensus" 11767..11782 /note="L2 repeat: matches 2692..2708 of consensus" 11783..12083 /note="Alusx repeat: matches 1..300 of consensus" 12084..12121 /note="L2 repeat: matches 2661..2692 of consensus"

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 01:41:55 ; Search time 503 Seconds  
(without alignments)  
9991.289 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccaggacc.....tctcttgatgatgctag 1183

Scoring table: OLIGO\_NUC  
Gapop 50.0, Gapext 60.0

Searched: 337863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002as.\*
- 7: Geneseq2003as.\*
- 8: Geneseq2003bs.\*
- 9: Geneseq2003cs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1183	100.0	1183	6	ABK61465 Human CDN
2	1132	95.7	2567	6	AAL43980 Human Src
3	735	62.1	786	6	AAL44089 Human mod
4	735	62.1	786	6	ABQ74343 Human Src
5	657	55.5	763	6	ABQ98670 Human ORF
6	592	50.0	837	3	AAC77202 Human ORF
7	483	40.8	737	6	AAL44090 Mouse MAR
8	468	39.6	1413	6	ABQ99374 Human cod
9	405	34.2	603	5	AS74748 DNA encod
10	401	33.9	2049	5	AS74750 DNA encod
11	348	23.4	444	6	ABQ98669 Human ORF
12	341	28.8	875	6	ABQ99151 Human ORF
13	244	20.6	445	5	AS74747 DNA encod
14	141	11.9	211	5	AS70181 DNA encod
15	134	11.3	432	4	AAL12879 Probe #28
16	134	11.3	432	4	ABA54580 Human foe
17	134	11.3	432	4	AAL134236 Probe #29
18	134	11.3	432	4	ABA44128 Human bre
19	134	11.3	432	4	ABA24363 Probe #28
20	134	11.3	432	4	AAL28314 Human bon
21	134	11.3	432	4	AAK02872 Human bra
22	134	11.3	432	4	ABS27912 Human liv
23	134	11.3	432	5	AAI02797 Probe #27

24	134	11.3	432	6	ABS02823	Human gen
25	134	11.3	448	4	AAI14520	Probe #44
26	134	11.3	448	6	ABS04499	Human gen
27	96	8.1	96	4	AAI22119	Probe #12
28	96	8.1	96	4	ABA67198	Human foe
29	96	8.1	96	4	AAI47414	Probe #16
30	96	8.1	96	4	ABA49284	Human bre
31	96	8.1	96	4	ABA4292	Probe #12
32	96	8.1	96	4	AAK41374	Human bon
33	96	8.1	96	4	AAK15640	Human bra
34	96	8.1	96	4	ABS40966	Human liv
35	96	8.1	96	5	AAI07818	Probe #78
36	96	8.1	96	6	ABS15380	Human gen
37	43	3.6	1348	6	AAL44087	Mouse mod
38	30	2.5	30	6	AAD43983	Human leu
39	26	2.2	26	6	ABK61506	Human NOY
40	25	2.1	25	6	AAL44100	Human mod
41	25	2.1	25	6	AAL44098	Human mod
42	25	2.1	25	6	AAL44099	Human mod
43	23	1.9	23	6	AAL44097	Human mod
44	22	1.9	22	6	ABK61507	Human NOV
45	22	1.9	3070	2	AAK55277	Nucleotid

## ALIGNMENTS

RESULT 1  
ABK61465  
ID ABK61465 standard; cDNA; 1183 BP.  
XX  
AC ABK61465;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human cDNA encoding protein NOV13.  
XX  
KW Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;  
KW cell signal processing disorder; metabolic pathway modulation disorder;  
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;  
KW uterus cancer; immune response; graft-versus-host disease;  
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;  
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;  
KW Albright hereditary osteodystrophy.  
XX  
OS Homo sapiens.  
XX  
FN WO200216599-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 27-AUG-2001; 2001WO-US026510.  
XX  
PR 25-AUG-2000; 2000US-0228191P.  
PR 08-FEB-2001; 2001US-0267300P.  
PR 20-FEB-2001; 2001US-026961P.  
PR 20-MAR-2001; 2001US-0277337P.  
XX  
PA (CURA-) CURAGEN CORP.  
PA (CORT-) COR THERAPEUTICS INC.  
XX  
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;  
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;  
XX  
DR WPI; 2002-280937/32.  
DR P-26DB; AAU91308.  
XX  
PT New polypeptides for treating or preventing a disorder associated with  
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.  
XX  
PS Claim 1; Page 98; 263pp; English.  
XX  
CC The invention relates to an isolated polypeptide (NOVX) a mature form of

CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide  
 CC encoding NOVX (or its complement, fragment or variant). NOVX is NOVX-14,  
 CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it,  
 CC and antibody against it, are useful for treating or preventing (e.g. by  
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOVX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation of Albrecht  
 CC hereditary osteodystrophy and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC for developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence encodes a NOVX protein.  
 XX  
 SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Query Match 100.0%; Score 1183; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AGCTAGAGCTCCAGAGACCCAGCGCTGTCTCTGTGACAGAGCTCAAGAGGCGCTGGG	60
Db	1	AGCTAGAGCTCCAGAGACCCAGCGCTGTCTCTGTGACAGAGCTCAAGAGGCGCTGGG	60
Qy	61	CTTTCCTCCCTGGCTGGCTGTCTGGAGAGGTTCCCGAGTCCAGAAATCCCTAAGGAG	120
Db	61	CTTTCCTCCCTGGCTGGCTGTCTGGAGAGGTTCCCGAGTCCAGAAATCCCTAAGGAG	120
Qy	121	CATGGGCGAGCTGATCCATCCCTGGTGTACAAACTGCTGACTGCAGACAGATGCTGAGCT	180
Db	121	CATGGGCGAGCTGATCCATCCCTGGTGTACAAACTGCTGACTGCAGACAGATGCTGAGCT	180
Qy	181	ACCCAAACACACCTAGCTCCCTCCCTGAGATCTCCAGGCTGAGAGAGTTCTGGGTG	240
Db	181	ACCCAAACACACCTAGCTCCCTCCCTGAGATCTCCAGGCTGAGAGAGTTCTGGGTG	240
Qy	241	TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTAACCTGTCCA	300
Db	241	TCCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTAACCTGTCCA	300
Qy	301	GCCAGAGCTGCTGCTCAGAGAGCTGTCTTCCAGAGCTTTGATGACAAACCAATTTCC	360
Db	301	GCCAGAGCTGCTGCTCAGAGAGCTGTCTTCCAGAGCTTTGATGACAAACCAATTTCC	360
Qy	361	CTCGATGATGTCTTCTGAGTGTCTCTGAGGACCAATGGGAAGTCTGCCAGCAGAAG	420
Db	361	CTCGATGATGTCTTCTGAGTGTCTCTGAGGACCAATGGGAAGTCTGCCAGCAGAAG	420
Qy	421	AAATATCTGCGAGCCAGCTGAGTCTCTGTCGACAGCCAGGACCTGTGACCAT	480
Db	421	AAATATCTGCGAGCCAGCTGAGTCTCTGTCGACAGCCAGGACCTGTGACCAT	480
Qy	481	GGAAGCAGAGAGAGCAGCCGCTGAGGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGG	540
Db	481	GGAAGCAGAGAGAGCAGCCGCTGAGGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGG	540
Qy	541	GGCGGAGCTGCTGAGATCTGGGAGCATTGACCTGCTCTGAGGATGGAGACTG	600
Db	541	GGCGGAGCTGCTGAGATCTGGGAGCATTGACCTGCTCTGAGGATGGAGACTG	600
Qy	601	GTGAGCGGTGCTGTCTGAAGTCTCAGCAGAGAGTATAATCCCGAGCGTCCACGTGGG	660
Db	601	GTGAGCGGTGCTGTCTGAAGTCTCAGCAGAGAGTATAATCCCGAGCGTCCACGTGGG	660

Qy	661	CAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGAACTGCT	720
Db	661	CAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGAAAGCAGAGGAACTGCT	720
Qy	721	GTGTGTACCTGGACACCTGGAGGGCTTCTTATCCGGAGAGCCAGACCCAGAGAGG	780
Db	721	GTGTGTACCTGGACACCTGGAGGGCTTCTTATCCGGAGAGCCAGACCCAGAGAGG	780
Qy	781	CTCTTACTCTGTGAGTCCGCTCAGCCGCTGATCTCTGGGACCGGATCAGACACTA	840
Db	781	CTCTTACTCTGTGAGTCCGCTCAGCCGCTGATCTCTGGGACCGGATCAGACACTA	840
Qy	841	CAGGATCCACTCCCTGTGACATGGCTGTGATCTTCATCTCCGGAGCCAGACCCAGAG	900
Db	841	CAGGATCCACTCCCTGTGACATGGCTGTGATCTTCATCTCCGGAGCCAGACCCAGAG	900
Qy	901	ACTCCAGGCGCTGGTGGACCACTTACTCTGAGTGGCGGATGACATCTGCTCCCTACTCA	960
Db	901	ACTCCAGGCGCTGGTGGACCACTTACTCTGAGTGGCGGATGACATCTGCTCCCTACTCA	960
Qy	961	GGAGCGCTGTGCTGAGAGGGCTGGCGCTCCCTGGCAAGGATATACCCCTACCTGT	1020
Db	961	GGAGCGCTGTGCTGAGAGGGCTGGCGCTCCCTGGCAAGGATATACCCCTACCTGT	1020
Qy	1021	GACTGTGAGAGGACACCACTCACTGGAAAGAGTGGAGCTGCGGATGACATCTGCTCTCTGA	1080
Db	1021	GACTGTGAGAGGACACCACTCACTGGAAAGAGTGGAGCTGCGGATGACATCTGCTCTCTGA	1080
Qy	1081	AGTGTGCCAGAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTA	1140
Db	1081	AGTGTGCCAGAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTA	1140
Qy	1141	CATCAGCTGAAATGACAGAGGCTCTCTTTGGATGATGCTCTAG 1183	
Db	1141	CATCAGCTGAAATGACAGAGGCTCTCTTTGGATGATGCTCTAG 1183	

RESULT 2

AAD43980  
 ID AAD43980 standard; cDNA; 2567 BP.

XX AAD43980;

DT 13-DEC-2002 (first entry)

XX Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.

Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 multiple sclerosis; asthma; acute respiratory distress syndrome;  
 pulmonary disorder; dermatological; neuroprotective; gene; ss.

OS Homo sapiens.

Key	Location/Qualifiers
CDS	415..1200
	/*tag= a
	/product= "Human SLAP-2"

WO200242457-A1.

PD 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US043367.

XX 22-NOV-2000; 2000US-0252545P.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;



30-MAY-2002.  
26-NOV-2001; 2001WO-CA001662.  
27-NOV-2000; 2000CA-02324663.  
(HOSP-) HOSPITAL FOR SICK CHILDREN.  
McGlade JC, Loreto MP;  
WPI: 2002-566564/60.  
P-PSDB; A015457.  
New isolated modulator of antigen receptor signaling protein or its fragment, useful for treating malignant disorders such as myeloid malignancies, autoimmune disorders and myeloproliferative disorders.  
Claim 12; Page 75; 110pp; English.  
The invention comprises the amino acid and coding sequences of modulator of antigen receptor signalling (MARS) proteins. The MARS protein is a putative tumour suppressor gene and exhibits structural and sequence similarity to the Src-like adaptor protein (SLAP). The MARS DNA and protein sequences of the invention are useful for the treatment of myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune disorders, immunosuppression, myeloproliferative disorders and malignancies related to the de-regulation of tyrosine kinases (e.g. breast cancer). The present cDNA sequence encodes a human MARS protein  
Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;  
Query Match 62.1%; Score 735; DB 6; Length 786;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
398 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCAAGCCCAAGCTTGATCTCTCTGTC 457  
1 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCAAGCCCAAGCTTGATCTCTCTGTC 60  
458 CAAAGCCAGGACCTGTGACCATGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 517  
61 CAAAGCCAGGACCTGTGACCATGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAT 120  
518 GGCAGTTTCCCGGAGAGTGGCCCGGAGCTGTGCTGAGACTCGGGAGGCAATTGACC 577  
121 GGCAGTTTCCCGGAGAGTGGCCCGGAGCTGTGCTGAGACTCGGGAGGCAATTGACC 180  
578 ATCGTCTCTGAGATGGAGATGGTGGACGCTGTCTGAGTCTCAGCAGAGAGAT 637  
181 ATCGTCTCTGAGATGGAGATGGTGGACGCTGTCTGAGTCTCAGCAGAGAGAT 240  
638 AACATCCCGCAGGCTCCAGTGGGAAAGTCTCCATGGGTGGCTGTATGAGGCGCTGAGC 697  
241 AACATCCCGCAGGCTCCAGTGGGAAAGTCTCCATGGGTGGCTGTATGAGGCGCTGAGC 300  
698 AGGAGAGAGAGAGAGAGTCTGTGTGTTTACTCGGGAACCTGAGGGGCTTCTCTCATC 757  
301 AGGAGAGAGAGAGAGTCTGTGTGTTTACTCGGGAACCTGAGGGGCTTCTCTCATC 360  
758 CGGAGAGAGAGAGAGTCTGTGTGTTTACTCGGGAACCTGAGGGGCTTCTCTCATC 817  
361 CGGAGAGAGAGAGAGTCTGTGTGTTTACTCGGGAACCTGAGGGGCTTCTCTCATC 420  
818 TCTGGGACCGGATCAGACACTACAGATCCACTGCGCTTGACATGGCTGGCTGTACATC 877  
421 TCTGGGACCGGATCAGACACTACAGATCCACTGCGCTTGACATGGCTGGCTGTACATC 480  
878 TCACGCGGCTCAGCTTCCCTCCTACTCCAGCCCTGTGTGACCATCTACTGAGCTGGG 937  
481 TCACGCGGCTCAGCTTCCCTCCTACTCCAGCCCTGTGTGACCATCTACTGAGCTGGG 540  
938 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGTCTCTGAGAGGGGCTGCCGCTCCCT 997

541 GATGACATCTGCTGCTACTCAAGAGCCCTGTCTCTGAGAGGCTGSCCGCTCCCT 600  
998 GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCACTGGAAGAGCTG 1057  
601 GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCACTGGAAGAGCTG 660  
1058 GACAGTCCCTCTCTGTTTCTGAAGTCCCAAGGGAGAGTCTCTCTCACTGAGGCT 1117  
661 GACAGTCCCTCTCTGTTTCTGAAGTCCCAAGGGAGAGTCTCTCTCACTGAGGCT 720  
1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATACAGGCTGCTCTTTGGATGAT 1177  
721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATACAGGCTGCTCTTTGGATGAT 780  
1178 GCTTAG 1183  
781 GCTTAG 786  
RESULT 4  
ABQ74343  
ID ABQ74343 standard; cDNA; 786 BP.  
XX AC ABQ74343;  
XX 15-OCT-2002 (first entry)  
XX Human Src-like inhibitory molecule (SLIM) encoding cDNA.  
XX Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;  
XX SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;  
XX modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;  
XX acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
XX chronic inflammatory disorder; autoimmune disorder; transplant rejection;  
XX gene; ss.  
XX Homo sapiens.  
XX OS  
XX Key Location/Qualifiers  
XX CDS 1..786  
XX /tag= a "SLIM"  
XX /product= "Src-like inhibitory molecule"  
XX W0200255707-A2.  
XX 18-JUL-2002.  
XX 10-JAN-2002; 2002WO-US000718.  
XX 10-JAN-2001; 2001US-0260953P.  
XX (RIGE-) RIGEL PHARM INC.  
XX Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
XX Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
XX WPI: 2002-575432/61.  
XX P-PSDB; ABP52187.  
XX New src-like inhibitory molecule protein, useful for treating  
XX immunodeficiency disorders and inflammatory disorders, comprises N-  
XX terminal myristylation sequence, SH2 domain and/or SH3 domain.  
XX Claim 3; Fig 2A; 91pp; English.  
XX The present sequence encodes the human Src-like inhibitory molecule  
XX (SLIM) protein (I). The present invention describes a SLIM protein  
XX comprising an N-terminal myristylation sequence, an N-terminal SH2  
XX domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising  
XX an N-terminal myristylation sequence and an N-terminal SH2 domain which  
XX is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and

[illegible]

KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
KW cancer; cardiovascular disease; allergy; autoimmune disease;  
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;

XX Homo sapiens.  
OS  
XX  
PN US2002082206-A1.

XX	.
PD	27-JUN-2002.
XX	
PF	30-MAY-2001; 2001US-00867550.

30-MAY-2000; 2000US-0208427P.  
XX  
PA (LEAC/) LEACH M D.

PA (METHUEN) MERRICK F.  
PA (CONL/) CONLEY P B.  
PA (TOPP/) TOPPER J N.  
PA (LAWD/) LAW D.

Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
WPI; 2002-626554/67.

New polypeptide designated ORFX are present in human atherogenic cells and are useful to prevent and treat ORFX-associated disorders including

PT. inflammatory disease.  
XX  
PS Claim 2; SEQ ID NO 953; 78pp; English.

CC The present invention relates to novel human ORFX polypeptides and their  
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ9267). The sequences  
CC were discovered in human atherogenic cells, in particular in platelets

many other tissues as well. Atherogenic cells are cells which have the potential to develop atherosclerotic plaques. The OREX polypeptides and nucleic acids are useful for treating or preventing a pathological

cardiovascular disease, allergy, autoimmune disease, wound healing, blood coagulation disorders or inflammatory disorders. Note: The sequence data for this patent did not form part of the printed specification, but was

CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
XX  
SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 55.5%; Score 657; DB 6; Length 763;  
Best Local Similarity 99.7%; Pred. No. 5.6e-300;  
Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0

117 GGAGCATGGGGCAGCTGATCCATCCCTGGTGTACAACTGCTGACTGCAGACAGATGCTG 178

1





immune deficiency; immune disorder; bacterial infection; allergy; cancer; fungal infection; autoimmune disorder; coagulation disorder; neutropenic; allergic; anti-inflammatory; immunosuppressive; neuroprotective; cytostatic; haemostatic; virucide; antibacterial; fungicide; immunostimulant; cerebroprotective; gene therapy; gene, ss.

OS Homo sapiens.  
XX WO200259250-A2.  
XX 01-AUG-2002.  
XX 16-NOV-2001; 2001WO-US042950.  
XX 17-NOV-2000; 2000US-00714936.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
PI Ren Z, Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX WPI; 2002-590824/63.  
DR N-PSDB; ABP64788.  
XX New isolated polynucleotide, useful in research, diagnostic or  
PT therapeutic methods, e.g. preventing or treating disorders involving  
PT aberrant protein expression or biological activity.  
XX  
XX Claim 1; SEQ ID NO 107; 394pp; English.

The present invention relates to novel human coding sequences (ABQ99268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in therapeutic, diagnostic and research methods. The polynucleotides may be used in the field of molecular biology as hybridisation probes, primers for PCR, for chromosome and gene mapping, for the recombinant production of protein, or in generation of anti-sense DNA or RNA. The polynucleotides are useful in diagnostics as expressed sequence tags (ESTs) for identifying expressed genes or for physical mapping of the human genome. The proteins may be used as molecular weight markers, or as nutritional sources or supplements. The proteins may be used to maintain and expand cell population in a totipotent or pluripotent state useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals or the development of bio-sensors. The polynucleotides and proteins are useful for preventing, treating or ameliorating disorders involving aberrant protein expression or biological activity, e.g. haematopoietic disorders, central/peripheral nervous system diseases, mechanical and traumatic disorders, non-healing wounds, immune deficiencies and disorders, infectious diseases caused by viral, bacterial or fungal infection, autoimmune disorders, allergic reactions and conditions, coagulation disorders, or cancer. The polynucleotide sequences of the invention were assembled from ESTs isolated mainly by sequencing by hybridisation, and in some cases, sequences obtained from one or more public databases. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;  
Query Match 39.6%; Score 468; DB 6; Length 1413;  
Best Local Similarity 99.6%; Pred No. 1e-210;  
Matches 568; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 345 TGACAAACCAATTCCTCGATGATGCTCTGAGTCTCTGCTGAGGAAACAAATGGGAA 404  
DB 1 TGACAAACCAATTCCTCGATGATGCTCTGAGTCTCTGCTGAGGAAACAAATGGGAG 60  
QY 405 GTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGCTCTGCTCAAGGCC 464  
DB 61 GCTGCCCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGCTCTGCTCAAGGCC 120  
QY 465 AGGACCTGTGACCATGAG 524

Db 121 AGGACCTGTGACCATGAG 180  
QY 525 TCCCGGACGCTGCGCCGCGGCGGAGCTGCTGAGACTCGGAGGCCATTGACCATCGTCT 584  
Db 181 TCCCGGACGCTGCGCCGCGGCGGAGCTGCTGAGACTCGGAGGCCATTGACCATCGTCT 240  
QY 585 CTGAGGATGAGAGACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 644  
Db 241 CTGAGGATGAGAGACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
QY 645 CCAGCGTCCACGCTGGGCAAGTCTCCCATGGGTGCTGATGAGGGCTGAGCAGGAGAGA 704  
Db 301 CCAGCGTCCACGCTGGGCAAGTCTCCCATGGGTGCTGATGAGGGCTGAGCAGGAGAGA 360  
QY 705 AACGAGAGAGACTGCTGAG 764  
Db 361 AACGAGAGAGACTGCTGAG 420  
QY 765 GCAGACACGAG 824  
Db 421 GCAGACACGAG 480  
QY 825 ACCGATCCAGACACTACAG 884  
Db 481 ACCGATCCAGACACTACAG 540  
QY 885 GCTCCTACCTTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 914  
Db 541 GCTCCTACCTTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 570

RESULT 9  
AAS74748  
ID AAS74748 standard; cDNA; 603 BP.  
XX  
AC AAS74748;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #10552.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
FR 31-MAR-2000; 2000US-00540217.  
FR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
WPI; 2001-639362/73.  
P-PSDB; ABG10561.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.  
PS Claim 1; SEQ ID NO 10552; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pcr\_sequences

XX SQ Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;  
 Query Match 34.2%; Score 405; DB 5; Length 603;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-181;  
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 779 GGCTCTTACTCTGTCAGTCGGCTCAGCCCGCTGATCTCGGACCGGATCAGACAC 838  
 Db 199 GGCTCTTACTCTGTCAGTCGGCTCAGCCCGCTGATCTCGGACCGGATCAGACAC 258  
 QY 839 TACAGGATCCATGCTTGACAAATGCTGCTGTACATCTCACCGCGCTCACCTTCCCC 898  
 Db 259 TACAGGATCCATGCTTGACAAATGCTGCTGTACATCTCACCGCGCTCACCTTCCCC 318  
 QY 899 TCACCTCAGCCCTGTCGACCATCTACTGAGCTGGCGGTGATCATCTGCTGCTACTC 958  
 Db 319 TCACCTCAGCCCTGTCGACCATCTACTGAGCTGGCGGTGATCATCTGCTGCTACTC 378  
 QY 959 AAGAGCCCTGTGCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1018  
 Db 379 AAGAGCCCTGTGCTCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438  
 QY 1019 GTGACTGTGAGAGGACACCACTCACTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 1078  
 Db 439 GTGACTGTGAGAGGACACCACTCACTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 498  
 QY 1079 GAAGCTGCCACAGGAGGAGTCTTCTTCTCAGTGGGTCTCCGGAGTCCCTCAGCTTC 1138  
 Db 499 GAAGCTGCCACAGGAGGAGTCTTCTTCTCAGTGGGTCTCCGGAGTCCCTCAGCTTC 558  
 QY 1139 TACATCAGCTGAATCAGCAGGCTGTCTCTTTGGATGATGCTAG 1183  
 Db 559 TACATCAGCTGAATCAGCAGGCTGTCTCTTTGGATGATGCTAG 603

RESULT 10  
 AAS74750  
 ID AAS74750 standard; cDNA; 2049 BP.  
 XX AAS74750;  
 AC AAS74750;  
 DT 13-FEB-2002 (first entry)  
 DX DNA encoding novel human diagnostic protein #10554.  
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 KW Homo sapiens.  
 OS W0200175067-A2.  
 PN W0200175067-A2.  
 XX 11-OCT-2001.  
 PD 30-MAR-2001; 2001WC-0008631.

XX 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYZE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-630362/73.  
 PI P-PSDB; ABG10563.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 10554; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pcr\_sequences

XX SQ Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;  
 Query Match 33.9%; Score 401; DB 5; Length 2049;  
 Best Local Similarity 99.8%; Pred. No. 4.6e-179;  
 Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 355 ATTTCCCTTCGATGATGCTTCTGAGTGTCTGCTGAGGACAAATGGAGTCTGCCAG 414  
 Db 922 ATTTCCCTTCGATGATGCTTCTGAGTGTCTGCTGAGGACAAATGGAGTCTGCCAG 981  
 QY 415 CAGAAGAAAATCTTGCAGAGCCCAAGCTTGAATTCCTCTGCTCCAGGCGGACCTGT 474  
 Db 982 CAGAAGAAAATCTTGCAGAGCCCAAGCTTGAATTCCTCTGCTCCAGGCGGACCTGT 1041  
 QY 475 GACCATGGAAGCAGAGAGCAAGCCACAGCCGTGGCCCTGGGCGAGTTTCCCGCAGG 534  
 Db 1042 GACCATGGAAGCAGAGAGCAAGCCACAGCCGTGGCCCTGGGCGAGTTTCCCGCAGG 1101  
 QY 535 TGGCCCGGCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCTCTCTGAGATGG 594  
 Db 1102 TGGCCCGGCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCTCTCTGAGATGG 1161  
 QY 595 AGACTGTGTGACCGGTGCTGTCTGAAGTCTCAGGACAGAGTATAACATCCCGAGGTCCA 654  
 Db 1162 AGACTGTGTGACCGGTGCTGTCTGAAGTCTCAGGACAGAGTATAACATCCCGAGGTCCA 1221  
 QY 655 CGTGGCAGAAAGTCTCCCATGGGTGGCTGTATGAGGCCCTGAGCAGGAGAGAGCAGAGA 714  
 Db 1222 CGTGGCAGAAAGTCTCCCATGGGTGGCTGTATGAGGCCCTGAGCAGGAGAGAGCAGAGA 1281  
 QY 715 ACTGCTGTGTATCTGGGAAACCTCGAGGGGCTTCTCTCATCCGGAGAGCAGACACAG 774  
 Db 1282 ACTGCTGTGTATCTGGGAAACCTCGAGGGGCTTCTCTCATCCGGAGAGCAGACACAG 1341

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:51 11 November 2014





GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 05:08:33 ; Search time 2929 Seconds  
(without alignments)  
12061.099 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccagagacc.....tctcttgatgatgctag i183

Scoring table: OLIGO NUC

Capop\_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_hci.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_hci.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pln.\*
- 20: em\_gss\_vrt.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_mam.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_phg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gss1.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	735	62.1	1201	9	AL541041
2	647	54.7	1002	12	BQ052308
3	578	48.9	986	12	BQ054265
4	568	48.0	1020	12	BQ054281

## ALIGNMENTS

RESULT 1  
AL541041  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CG08005YK23  
5-PRIME, mRNA sequence.  
AL541041  
EST.  
AL541041.2 GI:30544829  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (Bases 1 to 1201)  
Li, W.B.; Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:12871733.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 9825.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CG08005YK23&cluster=9825.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :

554	46.8	1069	12	BQ052468
459	38.8	616	13	BX383606
426	36.0	778	12	BG178487
405	34.2	566	12	BG284179
390	33.0	878	12	BQ053486
300	25.4	597	9	AL844311
276	23.3	614	9	AL844307
255	21.6	794	12	BG677567
205	17.3	960	13	BX944126
200	16.3	606	9	AL844309
116	9.8	611	9	AL844308
106	9.0	511	9	AL844310
103	8.7	506	28	B45150
80	6.8	639	9	AL844312
68	5.7	546	28	AQ556467
57	4.8	149	9	AA309769
47	4.0	553	28	AQ556478
43	3.6	660	10	BG35615
43	3.6	660	13	BY742155
43	3.6	926	11	AK020837
43	3.6	2637	11	AK088672
43	3.6	2974	11	AK030877
42	3.6	322	10	AW437301
42	3.6	569	14	CB426333
34	2.9	377	9	AA959151
30	2.5	389	10	BE015229
29	2.5	603	13	BQ553005
29	2.5	627	10	BG419854
27	2.3	757	29	CC527971
23	1.9	541	12	B1898999
22	1.9	784	12	BG388843
22	1.9	791	14	CB994934
22	1.9	875	12	BG282272
22	1.9	934	12	B1549288
22	1.9	1109	12	BM926459
22	1.9	2517	29	AY418068
22	1.8	2517	29	AY418069
21	1.8	320	14	CB557851
21	1.8	332	10	BE030537
21	1.8	363	9	AA365070
21	1.8	408	29	CG869927

BQ052468 AGENCOURT  
BX383606 BX383606  
BG178487 602328305  
BG284179 602408226  
BQ053486 AGENCOURT  
AL844311 AL844311  
AL844307 AL844307  
BG677567 6036224118  
BX944126 AGENCOURT  
AL844309 AL844309  
AL844308 AL844308  
AL844310 AL844310  
B45150 HS-1060-B1-  
AL844312 AL844312  
AQ556467 HS 5236 B  
AA309769 EST180699  
AQ556478 HS 5236 B  
BG35615 BG35615  
BY742155 BY742155  
AK020837 Mus muscu  
AK088672 Mus muscu  
AK030877 Mus muscu  
AW437301 78266 MAR  
CB426333 601508 MA  
AA959151 vz51g06.r  
BE015229 127457 MA  
BQ553005 H4019E02-  
BG419854 BG419854  
CC527971 CH240\_403  
B1898999 480839 MA  
BG388843 602414563  
CB994934 AGENCOURT  
BG282272 602402951  
B1549288 603189926  
BM926459 AGENCOURT  
AY418068 Homo sapi  
AY418069 Pan trogl  
CB557851 OSJNEC13C  
BE030537 128644 MA  
AA365070 EST15755  
CG869927 XS0853 SA

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DE005AF12QPI.

## FEATURES

Location/Qualifiers  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="CS0DE005YK23"  
/tissue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with NotI and cloned into  
the NotI and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

## ORIGIN

Query Match 62.1%; Score 735; DB 9; Length 1201;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 835; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TAGAGCTCCAGGACCCAGCGCTGTGTCTCTGTGACAGAGCTCAAGGGCCCTGGGCT 63  
Db 65 TAGAGCTCCAGGACCCAGCGCTGTGTCTGTGACAGAGCTCAAGGGCCCTGGGCT 124

Qy 64 TCCCTCCCTGGCTGGCTGTGTGTGGAGGGTTCCCGATCCAGATCCCTAAGAGCAT 123  
Db 125 TCCCTCCCTGGCTGGCTGTGTGGAGGGTTCCCGATCCAGATCCCTAAGAGCAT 184

Qy 124 GGGGAGCTGATCCATCCCTGGGTGTAACAACCTGCTGACTGCAGACAGATGCTGACTACC 183  
Db 185 GGGGAGCTGATCCATCCCTGGGTGTAACAACCTGCTGACTGCAGACAGATGCTGACTACC 244

Qy 184 CAACCAACACCTAGCTCTCCCTGAGATCTCCAGGCTGAGAGAGTTCTGGGTGTC 243  
Db 245 CAACCAACACCTAGCTCTCCCTGAGATCTCCAGGCTGAGAGAGTTCTGGGTGTC 304

Qy 244 TAGACCAAGACACTGGGAGACTTCAGAGGGGCCCCCAAGGCTCAACCTGTCAGCC 303  
Db 305 TAGACCAAGACACTGGGAGACTTCAGAGGGGCCCCCAAGGCTCAACCTGTCAGCC 364

Qy 304 AGAGCATGGCTCTCAGAGAGCTGTCTTCCAGAGCTTCCAGGCTGAGAGAGTTCTGGGTGTC 363  
Db 365 AGAGCATGGCTCTCAGAGAGCTGTCTTCCAGAGCTTCCAGGCTGAGAGAGTTCTGGGTGTC 424

Qy 364 GATGATGTCTTCTGAGTGTCTGCTGAGGAACTATGGGAAGTCTGCCCCAGAGAGAA 423  
Db 425 GATGATGTCTTCTGAGTGTCTGCTGAGGAACTATGGGAAGTCTGCCCCAGAGAGAA 484

Qy 424 ATCTCTCCAGCCCAAGCTTGATGTTCTCTGTCAGGCCAGGACCTGTGACCATGGA 483  
Db 485 ATCTCTCCAGCCCAAGCTTGATGTTCTCTGTCAGGCCAGGACCTGTGACCATGGA 544

Qy 484 AGCAGAGAGAGCAAGGCCCAAGCCCTGGCCCTGGGAGTTTCCCGGAGGTGGCCCGGC 543  
Db 545 AGCAGAGAGAGCAAGGCCCAAGCCCTGGCCCTGGGAGTTTCCCGGAGGTGGCCCGGC 604

Qy 544 CGAGCTCTCGCTGAGATCCGGAGAGCATGACCATCTGCTGTGAGAGTGAAGTGTG 603  
Db 605 CGAGCTCTCGCTGAGATCCGGAGAGCATGACCATCTGCTGTGAGAGTGAAGTGTG 664

Qy 604 GACGGTCTGTCTGAGTCTCAGGACAGAGTATACATCCCGAGCGTCCAGTGGGCAA 663  
Db 665 GACGGTCTGTCTGAGTCTCAGGACAGAGTATACATCCCGAGCGTCCAGTGGGCAA 724

Qy 664 AGTCTCCATGGGTGGTGTATCAGGGCTGAGCGGAGAGAGCAGAGGAGTCTGTT 723  
Db 725 AGTCTCCATGGGTGGTGTATAGGGCTGTGAGCGGAGAGAGCAGAGGAGTCTGTT 784

Qy 724 GTTACCTGGGAACCTTGGAGGGGCTTCTCTCATCCCGGAGAGCCAGACAGAGAGGCTC 783  
Db 785 GTTACCTGGGAACCTTGGAGGGGCTTCTCTCATCCCGGAGAGCCAGACAGAGAGGCTC 844

Qy

Db

## RESULT 2

BQ052308 1002 bp mRNA linear EST 29-MAR-2002  
LOCUS AGENCOURT 6868571 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5933542  
DEFINITION 5', mRNA sequence.

ACCESSION BQ052308  
VERSION BQ052308.1 GI:19811648

KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 1002)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaaps-remail.nih.gov

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LCM2118 row: d column: 23

High quality sequence stop: 670.

## FEATURES

## source

1..1002  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5933542"  
/tissue\_type="natural killer cells, cell line"  
/lab\_host="pH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_106"  
/note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 54.7%; Score 647; DB 12; Length 1002;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 747; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 333 CCAAGCCTTGATGACAAACCAATTCCTCGATGATGCTCTGAGTGTCTGCTGCTGAG 392

Db 31 CCAAGCCTTGATGACAAACCAATTCCTCGATGATGCTCTGAGTGTCTGCTGCTGAG 90

Qy 393 GAACAATGGAAAGTCTGCCAGAGAGAAAATCTTGCCAAAGCCCAAGCTTGAGTTCT 452

Db 91 GAACAATGGAAAGTCTGCCAGAGAGAAAATCTTGCCAAAGCCCAAGCTTGAGTTCT 150

Qy 453 CTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGTGG 512

Db 151 CTGTCCAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGCCACAGCCGTGG 210

Qy 513 CCTGGGAGTTTCCCGGAGTGGCCCGCGAGTGTCTGCTGAGACTCGGGAGGCAT 572

Db 211 CCTGGGAGTTTCCCGGAGTGGCCCGCGAGTGTCTGCTGAGACTCGGGAGGCAT 270

Qy 573 TGACCATCTCTCTCAGGATGGAGACTGGTGGACGGTGTCTGTGAAAGTCTCAGGACAG 632

Db 271 TGACCATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGTGAAGTCTCAGGCAGAG 330  
QY 633 AGTATAACATCCACAGCGTCCACCTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCC 692  
Db 331 AGTATAACATCCACAGCGTCCACCTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCC 390  
QY 693 TGACGAGGAGAAAGCAGAGAACTGCTGTGTGTACTCTGGAAACCTTGAGGGGCTTCC 752  
Db 391 TGAGCAGGAGAAAGCAGAGAACTGCTGTGTGTACTCTGGAAACCTTGAGGGGCTTCC 450  
QY 753 TCATCCGGAGAGCCAGACAGGAGAGGCTTCACTCTCTGTGAGTCCCGCTCAGCGGCC 812  
Db 451 TCATCCGGAGAGCCAGACAGGAGAGGCTTCACTCTCTGTGAGTCCCGCTCAGCGGCC 510  
QY 813 CTGATCTGGGCGGATCAGACACTACAGATCCACTCCCTTGACATGGCTGCTGT 872  
Db 511 CTGATCTGGGCGGATCAGACACTACAGATCCACTCCCTTGACATGGCTGCTGT 570  
QY 873 ACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCGCTGGTGGACCACTTCTGAGC 932  
Db 571 ACATCTCACCGCGCTCACCTTCCCTCACTCCAGGCGCTGGTGGACCACTTCTGAGC 630  
QY 933 TGGGGATGATCTGCTCTCAAGGAGCGCTGTCTCTGAGAGGCTGGCGGC 992  
Db 631 TGGGGATGATCTGCTCTCAAGGAGCGCTGTCTCTGAGAGGCTGGCGGC 690  
QY 993 TCCTTGGCAAGGATATACCTTCACTCTGAGTGGCAGAGGACCACTCAACTGGAAG 1052  
Db 691 TCCTTGGCAAGGATATACCTTCACTCTGAGTGGCAGAGGACCACTCAACTGGAAG 750  
QY 1053 AGCTGGACAGCTCCCTCTCTTCTGAA 1081  
Db 751 AGCTGGACAGCTCCCTCTCTTCTGAA 779

RESULT 3  
BQ054265  
LOCUS  
DEFINITION AGNCOURT 6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
5', mRNA Sequence.

ACCESSION BQ054265  
VERSION BQ054265.1 GI:19813605  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2125 row: 1 column: 12  
High quality sequence stop: 515.

FEATURES  
Location/Qualifiers  
1..986  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5936339"  
/tissue types="natural killer cells, cell line"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH MGC 106"  
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 48.9%; Score 578; DB 12; Length 986;  
Best Local Similarity 99.8%; Pred. No. 1.2e-281;  
Matches 628; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 120 GCATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGAGTGCAGACAGATGCTGAGC 179  
Db 1 GCATGGGGAGCTGATCCATCCCTGGTGTACAACTGCTGAGTGCAGACAGATGCTGAGC 60  
QY 180 TACCAAAACCAACACCTAGCCTTCCCTGAAGATCCCTCCAGGCTGAGAGAGTTCCTGGGT 239  
Db 61 TACCAAAACCAACACCTAGCCTTCCCTGAAGATCCCTCCAGGCTGAGAGAGTTCCTGGGT 120  
QY 240 GTCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTCC 299  
Db 121 GTCTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACCTGTCC 180  
QY 300 AGCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAATTTC 359  
Db 181 AGCCAGAGCATGCTCTCAGCAGAGCTGTCTTCCAAAGCCTTTGATGACAAACCAATTTC 240  
QY 360 CCTCGATGATGCTTCTGAGTGTCTGCTGAGGAGCAATGGGAAGTCTGCCACAGAA 419  
Db 241 CCTCGATGATGCTTCTGAGTGTCTGCTGAGGAGCAATGGGAAGTCTGCCACAGAA 300  
QY 420 GAAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTGTCCAGGCCAGGAGACTGTGACCA 479  
Db 301 GAAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTGTCCAGGCCAGGAGACTGTGACCA 360  
QY 480 TGAAGCAGAGAGAAACCAAGCCCAAGCTTGAAGTCTCTGTCCAGGCCAGGAGACTGTGACCA 539  
Db 361 TGAAGCAGAGAGAAACCAAGCCCAAGCTTGAAGTCTCTGTCCAGGCCAGGAGACTGTGACCA 420  
QY 540 CGGCCGAGTGTGCTGAGACTCGGGAGCCATTGACCAATCGTCTCTGAGGATGAGACT 599  
Db 421 CGGCCGAGTGTGCTGAGACTCGGGAGCCATTGACCAATCGTCTCTGAGGATGAGACT 480  
QY 600 GGTGGACGGTGTGCTGAGACTCGGGAGCCATTGACCAATCGTCTCTGAGGATGAGACT 659  
Db 481 GGTGGACGGTGTGCTGAGACTCGGGAGCCATTGACCAATCGTCTCTGAGGATGAGACT 540  
QY 660 GCAAGTCTCCATGGTGGCTGTATGAGGCGCTGAGCAGGGAGAAAGCAGAGGAAGTGC 719  
Db 541 CCAAGTCTCCATGGTGGCTGTATGAGGCGCTGAGCAGGGAGAAAGCAGAGGAAGTGC 600  
QY 720 TGTGTGTACTCTGGGAACCTCTGGAGGGGCC 748  
Db 601 TGTGTGTACTCTGGGAACCTCTGGAGGGGCC 629

## RESULT 4

BQ054281

LOCUS

DEFINITION

AGNCOURT 6830234 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936362

5', mRNA Sequence.

ACCESSION

BQ054281

VERSION

BQ054281.1 GI:19813621

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

QY	791	CTGTTCAGTCCCGCCTCAGCGCGCCTGCATCTTGGGACCGGATCAGACACTACAGGATCCAC	850
Db	605	CTGTTCAGTCCCGCCTCAGCGCGCCTGCATCTTGGGACCGGATCAGACACTACAGGATCCAC	664
QY	851	TGCCCTTGACATGGCTGGCTGTACATCTCACCGCGCCTCACCTTCCCTCTCACTCCAGGCC	910
Db	665	TGCCCTTGACATGGCTGGCTGTACATCTCACCGCGCCTCACCTTCCCTCTCACTCCAGGCC	724
QY	911	C 911	
Db	725	C 725	

  

RESULT 5	BQ052468	
LOCUS	AGENCOURT_6869422_NIH_MGC_106	Homo sapiens
DEFINITION	linear	EST 29-MAR-2002
ACCESSION	BQ052468	
VERSION	BQ052468.1	GI:19811808
KEYWORDS	5', mRNA sequence.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 1069)	
	NIH-MGC <a href="http://imgc.nci.nih.gov/">http://imgc.nci.nih.gov/</a> .	

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
<b>Accession:</b>	
<b>Accession number:</b>	
<b>Accession type:</b>	
<b>Accession year:</b>	
<b>Accession month:</b>	
<b>Accession day:</b>	
<b>Accession date:</b>	
<b>Accession time:</b>	
<b>Accession location:</b>	
<b>Accession status:</b>	
<b>Accession notes:</b>	
<b>Accession comments:</b>	
<b>Accession history:</b>	
<b>Accession details:</b>	
<b>Accession summary:</b>	
<b>Accession description:</b>	
<b>Accession information:</b>	
<b>Accession metadata:</b>	
<b>Accession data:</b>	
<b>Accession file:</b>	
<b>Accession folder:</b>	
<b>Accession path:</b>	
<b>Accession url:</b>	
<b>Accession link:</b>	
<b>Accession id:</b>	
<b>Accession key:</b>	
<b>Accession value:</b>	
<b>Accession name:</b>	
<b>Accession title:</b>	
<b>Accession subtitle:</b>	
<b>Accession author:</b>	
<b>Accession editor:</b>	
<b>Accession publisher:</b>	
<b>Accession distributor:</b>	
<b>Accession agent:</b>	
<b>Accession contact:</b>	
<b>Accession address:</b>	
<b>Accession phone:</b>	
<b>Accession fax:</b>	
<b>Accession email:</b>	
<b>Accession website:</b>	
<b>Accession homepage:</b>	
<b>Accession portal:</b>	
<b>Accession gateway:</b>	
<b>Accession interface:</b>	
<b>Accession application:</b>	
<b>Accession service:</b>	
<b>Accession tool:</b>	
<b>Accession software:</b>	
<b>Accession hardware:</b>	
<b>Accession system:</b>	
<b>Accession platform:</b>	
<b>Accession architecture:</b>	
<b>Accession design:</b>	
<b>Accession implementation:</b>	
<b>Accession deployment:</b>	
<b>Accession operation:</b>	
<b>Accession maintenance:</b>	
<b>Accession support:</b>	
<b>Accession training:</b>	
<b>Accession documentation:</b>	
<b>Accession help:</b>	
<b>Accession manual:</b>	
<b>Accession guide:</b>	
<b>Accession tutorial:</b>	
<b>Accession example:</b>	
<b>Accession demo:</b>	
<b>Accession test:</b>	
<b>Accession benchmark:</b>	
<b>Accession performance:</b>	
<b>Accession speed:</b>	
<b>Accession accuracy:</b>	
<b>Accession reliability:</b>	
<b>Accession security:</b>	
<b>Accession privacy:</b>	
<b>Accession integrity:</b>	
<b>Accession availability:</b>	
<b>Accession accessibility:</b>	
<b>Accession interoperability:</b>	
<b>Accession compatibility:</b>	
<b>Accession portability:</b>	
<b>Accession scalability:</b>	
<b>Accession flexibility:</b>	
<b>Accession extensibility:</b>	
<b>Accession modifiability:</b>	
<b>Accession reusability:</b>	
<b>Accession sustainability:</b>	
<b>Accession longevity:</b>	
<b>Accession resilience:</b>	
<b>Accession robustness:</b>	
<b>Accession fault tolerance:</b>	
<b>Accession recovery:</b>	
<b>Accession backup:</b>	
<b>Accession restore:</b>	
<b>Accession update:</b>	
<b>Accession patch:</b>	
<b>Accession upgrade:</b>	
<b>Accession downgrade:</b>	
<b>Accession rollback:</b>	
<b>Accession revert:</b>	
<b>Accession reset:</b>	
<b>Accession restart:</b>	
<b>Accession shutdown:</b>	
<b>Accession power off:</b>	
<b>Accession power on:</b>	
<b>Accession boot:</b>	
<b>Accession login:</b>	
<b>Accession logout:</b>	
<b>Accession session:</b>	
<b>Accession user:</b>	
<b>Accession role:</b>	
<b>Accession permission:</b>	
<b>Accession privilege:</b>	
<b>Accession right:</b>	
<b>Accession authority:</b>	
<b>Accession jurisdiction:</b>	
<b>Accession domain:</b>	
<b>Accession scope:</b>	
<b>Accession range:</b>	
<b>Accession extent:</b>	
<b>Accession coverage:</b>	
<b>Accession reach:</b>	
<b>Accession impact:</b>	
<b>Accession influence:</b>	
<b>Accession effect:</b>	
<b>Accession result:</b>	
<b>Accession outcome:</b>	
<b>Accession consequence:</b>	
<b>Accession implication:</b>	
<b>Accession significance:</b>	
<b>Accession importance:</b>	
<b>Accession relevance:</b>	
<b>Accession applicability:</b>	
<b>Accession suitability:</b>	
<b>Accession appropriateness:</b>	
<b>Accession feasibility:</b>	
<b>Accession practicality:</b>	
<b>Accession viability:</b>	
<b>Accession possibility:</b>	
<b>Accession probability:</b>	
<b>Accession likelihood:</b>	
<b>Accession chance:</b>	
<b>Accession risk:</b>	
<b>Accession hazard:</b>	
<b>Accession danger:</b>	
<b>Accession threat:</b>	
<b>Accession peril:</b>	
<b>Accession jeopardy:</b>	
<b>Accession predicament:</b>	
<b>Accession dilemma:</b>	
<b>Accession quandary:</b>	
<b>Accession conundrum:</b>	
<b>Accession puzzle:</b>	
<b>Accession mystery:</b>	
<b>Accession enigma:</b>	
<b>Accession riddle:</b>	
<b>Accession paradox:</b>	
<b>Accession contradiction:</b>	
<b>Accession inconsistency:</b>	
<b>Accession incompatibility:</b>	
<b>Accession irreconcilable:</b>	

COMMENT	ORIGIN
Contact: Robert Strausberg, Ph.D. Email: cga@rsb@mail.nih.gov Tissue Procurement: Dr. Daniel McVicar, DBS/NCI CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2118 row: n column: 13 High quality sequence stop: 681. Location/Qualifiers 1. .1069 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5933772" /tissue_type="natural killer cells, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 106" /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."	Query Match 46.8%; Score 554; DB 12; Length 1069; Best Local Similarity 99.8%; Pred. No. 1.9e-269; Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
	310 TGCCTCTCAGCAGAGCTGTCTCCCAAGCCCTTTGATGACAAACCAATTTCCCTCGATGAT 369
	1 TGCCTCTCAGCAGAGCTGTCTCCCAAGCCCTTTGATGACAAACCAATTTCCCTCGATGAT 60
	370 GTGCTTCTGAGTCTGTCTGCGAGCAATGCGGAAGTCTCCAGCAGACAGAAATCTCT 429
	61 GTGCTTCTGAGTCTGTCTGCGAGCAATGCGGAAGTCTCCAGCAGACAGAAATCTCT 120
	430 GCCAAGCCCCAAGCTTGATGTTCTCTGTGCCAAGGCCAGGGACCTGTGTACCATGGAAACAGA 489

ORIGIN

Db 121 GCCAAGCCCAAGCTTGTCTCTGTCTCAAGGCCAGGAGCTGTGACCATGGAAGCAGA 180  
Qy 490 GAGAGCAGGCCACAGCCGCTGGCCCTGGCAGATTTCGCCGAGGTGGCCGCGGAGCT 549  
Db 181 GAGAAGCAAGGCCACAGCCGCTGGCCCTGGCAGATTTCGCCGAGGTGGCCGCGGAGCT 240  
Qy 550 GTCCGTGAGACTCGGGGAGCCATTGACCATCGTCTCTGAGGATGAGACTGTGTGACCGT 609  
Db 241 GTCCGTGAGACTCGGGGAGCCATTGACCATCGTCTCTGAGGATGAGACTGTGTGACCGT 300  
Qy 610 GCTGTCTGAGTCTCAGCAGAGATATACATCCCGAGCTCCACGTGGGCAAGTCTC 669  
Db 301 GCTGTCTGAGTCTCAGCAGAGATATACATCCCGAGCTCCACGTGGGCAAGTCTC 360  
Qy 670 CCATGGTGGCTGTATGAGGGGCTGAGCAGGGAGAAAGCAGAGGAATCTGTGTGTACC 729  
Db 361 CCATGGTGGCTGTATGAGGGGCTGAGCAGGGAGAAAGCAGAGGAATCTGTGTGTACC 420  
Qy 730 TGGGACCTGAGGGGCTTCTCTATCCGGAGAGCCAGACGAGAGCTCTTACTC 789  
Db 421 TGGGAACCTGAGGGGCTTCTCTATCCGGAGAGCCAGACGAGAGCTCTTACTC 480  
Qy 790 TCTGTCACTCCGCTCAGCGGCTCTGATCTCTGGGACCGGATCAGACTACAGGATCCA 849  
Db 481 TCTGTCACTCCGCTCAGCGGCTCTGATCTCTGGGACCGGATCAGACTACAGGATCCA 540  
Qy 850 CTGCTTGAATATGGCTGGCTGTATCTCAGCGGCTCTCAGCTTCCCTCAGTCCAGGC 909  
Db 541 CTGCTTGAATATGGCTGGCTGTATCTCAGCGGCTCTCAGCTTCCCTCAGTCCAGGC 600  
Qy 910 CTTGG 914  
Db 601 CTTGG 605

RESULT 6  
BX383606  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BX383606 616 bp mRNA linear EST 08-MAY-2003  
BX383606 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
Homo sapiens cDNA clone CS0DJ013YK10 5-PRIME, mRNA sequence.  
BX383606  
EST.  
BX383606.1 GI:30457152  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 616)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 9825.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DJ013BF05QPI&cluster=9825.r. Contact :  
Peng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DJ013BF05QPI.  
Location/Qualifiers  
1. .516  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DJ013YK10"  
/cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/cell\_line="JURKAT"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT

10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 38.8%; Score 459; DB 13; Length 616;  
Best Local Similarity 99.8%; Pred. No. 2.7e-221;  
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTCACAGGACCCACGCCCTGTCTCTGTGACAGAGCTCAAGGGGCCCTGGGCTTCCCT 68  
Db 58 CTCACAGGACCCACGCCCTGTCTCTGTGACAGAGCTCAAGGGGCCCTGGGCTTCCCT 117  
Qy 69 CCTGGCTGGCTGTCTCTGGGAGGTTCCCAAGTCCAGATCCCTTAAGGAGATGGGGC 128  
Db 118 CCTGGCTGGCTGTCTGGGAGGTTCCCAAGTCCAGATCCCTTAAGGAGATGGGGC 177  
Qy 129 AGCTGATCCATCCCTGCTGTACAACTGCTGACTGCAGACAGATGCTGAGCTACCCAAAC 188  
Db 178 AGCTGATCCATCCCTGCTGTACAACTGCTGACTGCAGACAGATGCTGAGCTACCCAAAC 237  
Qy 189 CAACACTAGCCTCTCCCTGAAGATCTCCAGGCTGAGAGATTTGGGTGCTTAGGA 248  
Db 238 CAACACTAGCCTCTCCCTGAAGATCTCCAGGCTGAGAGATTTGGGTGCTTAGGA 297  
Qy 249 CCAAGGACACTGCAGACTTCCAGAGGGGCCCCCAAGCCCTAACCTGCCAGCCAGAGC 308  
Db 298 CCAAGGACACTGCAGACTTCCAGAGGGGCCCCCAAGCCCTAACCTGCCAGCCAGAGC 357  
Qy 309 ATGCGTCTCAGCAGAGCTGTCTTCCCAAGCTTTTGATGACAAACCAATTTCCCTCGATGA 368  
Db 358 ATGCGTCTCAGCAGAGCTGTCTTCCCAAGCTTTTGATGACAAACCAATTTCCCTCGATGA 417  
Qy 369 TGTGCTTCTGAGTCTCTGCTGAGGACATGGAAGTCTGCCAGCAGAGAAATCTC 428  
Db 418 TGTGCTTCTGAGTCTCTGCTGAGGACATGGAAGTCTGCCAGCAGAGAAATCTC 477  
Qy 429 TGCAAGCCCAAGCTTGAGTTCCTCTGTCCAAGGCCAGGACCTGTGACCATGGAAGCAG 488  
Db 478 TGCAAGCCCAAGCTTGAGTTCCTCTGTCCAAGGCCAGGACCTGTGACCATGGAAGCAG 537  
Qy 489 AGAGAACCAAGGCCACAGCCGTGGCCCTGG 518  
Db 538 AGAGAACCAAGGCCACAGCCGTGGCCCTGG 567

RESULT 7  
BX178487  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BX178487 778 bp mRNA linear EST 06-FEB-2001  
602328305F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4429896 5',  
mRNA sequence.  
BX178487  
EST.  
BX178487.1 GI:12685190  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 778)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: csaps@remail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LLAM10182 row: i column: 01  
High quality sequence stop: 657.  
Location/Qualifiers  
1..778

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:4429896"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_91"  
/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 36.0%; Score 426; DB 12; Length 778;  
Best Local Similarity 99.7%; Pred. No. 1.7e-204;  
Matches 596; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 62 CTTCCCTCCCTGGCTCGCTGTGTTGGAGGGTTCCCACTCCAGATCCCTAAGGAGC 121  
Db |  
QY 26 CTTCCCTCCCTGGCTCGCTGTGTTGGAGGGTTCCCACTCCAGATCCCTAAGGAGA 85  
Db |  
QY 122 ATGGGGCAGCTGATCCATCCCTGGTGTACAAACTGCTGACTGCAGACAGATGCTGAGCTA 181  
Db |  
QY 86 ATGGGGCAGCTGATCCATCCCTGGTGTACAAACTGCTGACTGCAGACAGATGCTGAGCTA 145  
QY 182 CCCAACCAACACCTAGCTCTCCCTGAGATCTCCCACTCCAGCTGAGAGAGTTCTGGGTGT 241  
Db |  
QY 146 CCCAACCAACACCTAGCTCTCCCTGAGATCTCCCACTCCAGCTGAGAGAGTTCTGGGTGT 205  
QY 242 CTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTTAACCTGTCCAG 301  
Db |  
QY 206 CTTAGGACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTTAACCTGTCCAG 265  
QY 302 CCAGAGCATGGCTTCCAGCAGAGCTGTCTCCAGAGCTTTGATGACAAACCAATTTCCC 361  
Db |  
QY 266 CCAGAGCATGGCTTCCAGCAGAGCTGTCTCCAGAGCTTTGATGACAAACCAATTTCCC 325  
QY 362 TCGATGATGTGCTCTGAGTGTCTGCTGAGGACCAATGGGAAGTGTGCCAGCAGAGA 421  
Db |  
QY 386 AAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCCCAAGGCCA-GGACCTGTGACCATG 444  
QY 482 GAAGCAGAGAGAGCAAGGCCACACCGCTGGCCCTGGGACGTTTCCCGCAGGTGGCCCG 541  
Db |  
QY 445 GAAGCAGAGAGAGCAAGGCCACACCGCTGGCCCTGGGACGTTTCCCGCAGGTGGCCCG 504  
QY 542 GCCGAGCTGTCTGAGACTCGGGAGCATTGACCATCGTCTCTGAGGATGGAGACTGG 601  
Db |  
QY 505 GCCGAGCTGTCTGAGACTCGGGAGCATTGACCATCGTCTCTGAGGATGGAGACTGG 564  
QY 602 TGGACGGTGTCTGAGTCTCAGGAGAGATATACATCCCAAGCTCCAGCTGG 659  
Db |  
QY 565 TGGACGGTGTCTGAGTCTCAGGAGAGATATACATCCCAAGCTCCAGCTGG 622

RESULT 8  
BG284179  
LOCUS  
DEFINITION  
602408226F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4520382 5',  
mRNA sequence.  
BG284179  
VERSION  
KEYWORDS  
EST.  
SOURCE  
ORGANISM  
Homo sapiens  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 566)  
NIH-MGC http://mgs.mcg.mcg.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: DCID/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10418 row: c column: 07  
High quality sequence start: 2  
High quality sequence stop: 566.  
Location/Qualifiers  
1..566

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:4520382"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_91"  
/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 34.2%; Score 405; DB 12; Length 566;  
Best Local Similarity 100.0%; Pred. No. 6.8e-134;  
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 779 GGCTTTACTCTCTGTCAGTCCGCTCAGCCGCTGATCTCCTGGACCGGATCAGACAC 838  
Db |  
QY 13 GGCTTTACTCTCTGTCAGTCCGCTCAGCCGCTGATCTCCTGGACCGGATCAGACAC 72  
Db |  
QY 839 TACAGGATCCATCGCTTGACAAATGGCTGGCTGTACATCTCACCAGCTCCTACCTTCCCC 898  
Db |  
QY 73 TACAGGATCCATCGCTTGACAAATGGCTGGCTGTACATCTCACCAGCTCCTACCTTCCCC 132  
QY 899 TCATCCAGGCGCTGGTGGACCAATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTC 958  
Db |  
QY 133 TCATCCAGGCGCTGGTGGACCAATTAATCTGAGCTGGCGGATGACATCTGCTGCTACTC 192  
QY 959 AAGGAGCCCTGTGCTGTCAGAGAGGCTGGCCGCTCCCTGGCAAGATATACCCCTACCT 1018  
Db |  
QY 193 AAGGAGCCCTGTGCTGTCAGAGAGGCTGGCCGCTCCCTGGCAAGATATACCCCTACCT 252  
QY 1019 GTGACTGTGTCAGAGGACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCT 1078  
Db |  
QY 253 GTGACTGTGTCAGAGGACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCT 312  
QY 1079 GAAGCTGCCACAGGAGGAGTCTTCTCTGAGGGTCTCCGAGAGTCCCTCAGCTTC 1138  
Db |  
QY 313 GAAGCTGCCACAGGAGGAGTCTTCTCTGAGGGTCTCCGAGAGTCCCTCAGCTTC 372  
QY 1139 TACATCAGCTCAATGACGAGGCTGTCTTTTGGATGATGCTAG 1183  
Db |  
QY 373 TACATCAGCTCAATGACGAGGCTGTCTTTTGGATGATGCTAG 417

RESULT 9  
BQ053486  
LOCUS  
DEFINITION  
5', mRNA sequence.  
BQ053486  
ACCESSION

BQ053486  
AGENCOURT\_6822017 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5935253  
5', mRNA sequence.  
BQ053486

BQ053486.1 GI:19812826  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 878)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LNCM2122 row: 1 column: 06  
 High quality sequence stop: 394.  
 FEATURES  
 source  
 1..878  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:595253"  
 /tissue\_type="natural killer killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC 106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC library."  
 ORIGIN  
 Query Match 33.0%; Score 390; DB 12; Length 878;  
 Best Local Similarity 99.8%; Pred. No. 3.3e-186;  
 Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 179 CTACCCAAACCAACCTAGCCTTCCTCGAAGATCTCCAGAGGCTTCTGGG 238  
 DB 54 CTACCCAAACCAACCTAGCCTTCCTCGAAGATCTCCAGAGGCTTCTGGG 113  
 QY 239 TGTCTTAGGACCAAGGACACTGGGAGACTTCCAGAAGGGCCCCAAAGCCCTTAACCTGTC 298  
 DB 114 TGTCTTAGGACCAAGGACACTGGGAGACTTCCAGAAGGGCCCCAAAGCCCTTAACCTGTC 173  
 QY 299 CAGCCAGAGCATGGGCTTCAGCAGAGCTGTCTTCCCAAGCGTTTGATGACAAACCAATTT 358  
 DB 174 CAGCCAGAGCATGGGCTTCAGCAGAGCTGTCTTCCCAAGCGTTTGATGACAAACCAATTT 233  
 QY 359 CCCTCGATGATGTCTTCTGAGTGCTCTGTGAGGAACAATGGGAAGCTTGCCTCCAGCAGA 418  
 DB 234 CCCTCGATGATGTCTTCTGAGTGCTCTGTGAGGAACAATGGGAAGCTTGCCTCCAGCAGA 293  
 QY 419 AGAAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTTCMAGGCCAAGGACCTGTGACC 478  
 DB 294 AGAAAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTTCMAGGCCAAGGACCTGTGACC 353  
 QY 479 ATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTGGGCACTTCCCGGCAAGGTGGC 538  
 DB 354 ATGGAAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTGGGCACTTCCCGGCAAGGTGGC 413  
 QY 539 CCGGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGAC 598  
 DB 414 CCGGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGGAGAC 473  
 QY 599 TGGTGGACGGTGTCTGTGAA 619

```

Db          474  TGGTGGACGGTCTGCTGTCAA 494
|||||
RESULT 10
LOCUS      AL844311                597 bp  mRNA  linear  EST 30-JUL-2002
DEFINITION AL844311 pool_YT_lib_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION  AL844311
VERSION     AL844311.1  GI:22019089
KEYWORDS   EST.
ORGANISM   Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 597)
AUTHORS    Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
            Sheridan,E.
TITLE      Homo sapiens EST sequence
JOURNAL    Unpublished (2002)
COMMENT    Contact: The Sanger Centre
            The Sanger Centre
            Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: humquery@sanger.ac.uk
            Sanger Centre name : scd01818.154136A
            Homo sapiens EST sequence. This sequence was generated as part of
            the Wellcome Trust Sanger Institute program to identify and
            annotate genes in the human genome. Incomplete or unconfirmed genes
            are experimentally analysed using a variety of cDNA library
            resources. This sequence was obtained from a PCR product generated
            from a pool of up to 100,000 cDNA clones derived from
            pool_YT_lib_v_SPD cDNA library. Further information can be found at
            http://www.sanger.ac.uk/Teams/Team69/.
FEATURES             Location/Qualifiers
     source            1..597
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /map="20"
                        /clone_lib="pool_YT_lib_v_SPD"
ORIGIN
Query Match      25.4%; Score 300; DB 9; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.4e-140;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884  CGCCTCACCTTCCCTCCTCAGTCCAGGGCCCTGGTGGACCATTACTCTGAGCTGGCGGATGAC 943
Db 37  CGCCTCACCTTCCCTCCTCAGTCCAGGGCCCTGGTGGACCATTACTCTGAGCTGGCGGATGAC 96
QY 944  ATCTGCTGCTACTCTCAGGAGCCCTGTCTCTCAGAGAGGCTGGCCCGCTCCCTGGCAAG 1003
Db 97  ATCTGCTGCTACTCTCAGGAGCCCTGTCTCTCAGAGAGGCTGGCCCGCTCCCTGGCAAG 156
QY 1004 GATATACCCCTACCTGTGACTGTGCAGAGACACCACTCAACTGGAAAGAGCTGGACAGC 1063
Db 157 GATATACCCCTACCTGTGACTGTGCAGAGACACCACTCAACTGGAAAGAGCTGGACAGC 216
QY 1064 TCCTCTCTGTTTCTGAGAGTGCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGG 1123
Db 217 TCCTCTCTGTTTCTGAGAGTGCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGG 276
QY 1124 GAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTCCTAG 1183
Db 277 GAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAGGCTGTCTCTTTGGATGATGCTCCTAG 336

RESULT 11
AL844307
LOCUS      AL844307                614 bp  mRNA  linear  EST 30-JUL-2002
DEFINITION AL844307 pool_AK_lib_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION  AL844307
VERSION     AL844307.1  GI:22019089

```

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 614)  
AUTHORS Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and Sheridan, E.  
TITLE Homo sapiens EST sequence  
JOURNAL Unpublished (2002)  
COMMENT Contact: The Sanger Centre  
The Sanger Centre  
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: humquerry@sanger.ac.uk  
Sanger Centre name : sccdi0816.154136A  
Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool\_AK lib v SPD cDNA library. Further information can be found at <http://www.sanger.ac.uk/Teams/Team69/>.

FEATURES  
source  
1..614  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="20"  
/clone\_lib="pool\_AK lib v SPD"

ORIGIN  
Query Match 23.3%; Score 276; DB 9; Length 614;  
Best Local Similarity 100.0%; Pred. No. 2.2e-128;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 908 GCCCTGGTGGACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCAAGAGGCC 967  
Db 44 GCCCTGGTGGACCACTTACTCTGAGCTGGCGGATGACATCTGCTGCTACTCAAGAGGCC 103  
QY 968 TGTGTCTTCGAGAGGCTGGCCCGCTCCCTGGGAGGATATACCCCTACTGTGACTGTG 1027  
Db 104 TGTGTCTTCGAGAGGCTGGCCCGCTCCCTGGGAGGATATACCCCTACTGTGACTGTG 163  
QY 1028 CAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAGCTGCC 1087  
Db 164 CAGAGGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAGCTGCC 223  
QY 1088 ACAGGGAGGAGTCTCTTCTCAGTGAGGCTTCGCGGAGTCCCTCAGCTTCTACATCAGC 1147  
Db 224 ACAGGGAGGAGTCTCTTCTCAGTGAGGCTTCGCGGAGTCCCTCAGCTTCTACATCAGC 283  
QY 1148 CTGATGACGAGGCTGCTCTTTGGATGATGCTAG 1183  
Db 284 CTGATGACGAGGCTGCTCTTTGGATGATGCTAG 319

RESULT 12  
BG677567  
LOCUS 794 bp mRNA linear EST 01-MAY-2001  
DEFINITION 602624118F1 NCI\_CGAP\_Skn4 Homo sapiens cDNA clone IMAGE:4748884 5', mRNA sequence.  
ACCESSION BG677567  
VERSION BG677567.1 GI:13908964  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 794)  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: James Cleaver, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLNL0601 row: d column: 05  
High quality sequence stop: 790.  
Location/Qualifiers  
1..794  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:4748884"  
/tissue\_type="squamous cell carcinoma"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_lib="NCI CGAP\_Skn4"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

ORIGIN  
Query Match 21.6%; Score 255; DB 12; Length 794;  
Best Local Similarity 100.0%; Pred. No. 1.1e-117;  
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 929 GAGCTGGCGGATGACATCTGCTGCTACTCAGAGGACCTGTGCTCGAGAGGCTGCG 988  
Db 17 GAGCTGGCGGATGACATCTGCTGCTACTCAGAGGACCTGTGCTCGAGAGGCTGCG 76  
QY 989 CGCTCCCTCCCGAAGATATACCCCTACTGTGACTGTGCGAGGACACCACTCAACTGG 1048  
Db 77 CGCTCCCTCCCGAAGATATACCCCTACTGTGACTGTGCGAGGACACCACTCAACTGG 136  
QY 1049 AAAGAGCTGGAGAGCTCCCTCCTGTTTCTGAAGTCCACAGGGAGGAGTCTTCTC 1108  
Db 137 AAAGAGCTGGAGAGCTCCCTCCTGTTTCTGAAGTCCACAGGGAGGAGTCTTCTC 196  
QY 1109 AGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACAGGCTGTCT 1168  
Db 197 AGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACAGGCTGTCT 256  
QY 1169 TTGGATGATGCTAG 1183  
Db 257 TTGGATGATGCTAG 271

RESULT 13  
BG677567  
LOCUS 960 bp mRNA linear EST 18-OCT-2002  
DEFINITION AGENCOURT\_10545003 NIH\_MGC\_107 Homo sapiens cDNA clone IMAGE:6728350 5', mRNA sequence.  
ACCESSION BG677567  
VERSION BG677567.1 GI:24132945  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 960)  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLCM3049 row: m column: 21  
High quality sequence stop: 628.

Location/Qualifiers

FEATURES

source

1..960  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6728350"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="PH108 (phage-resistant)"  
/clone\_lib="NIH MGC 107"  
/note="Organ: Breast; Vector: pOTB7; Site: 1: EcoRI; Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 17.3%; Score 205; DB 13; Length 960;  
Best Local Similarity 100.0%; Pred. No. 2.8e-92;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 979 GAGGGTGGCCCGCTCCCTGGGAGGATATACCTGATCTGATGTCGACAGACACC 1038  
Db 177 GAGGGTGGCCCGCTCCCTGGGAGGATATACCTGATCTGATGTCGACAGACACC 236  
  
Qy 1039 ACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGGA 1098  
Db 237 ACTCAACTGGAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGGA 296  
  
Qy 1099 GTCTCTTCTCAGTAGGGTCTCCGGAGTCCTCAGCTTCTACATCAGCTGAATGACGA 1158  
Db 297 GTCTCTTCTCAGTAGGGTCTCCGGAGTCCTCAGCTTCTACATCAGCTGAATGACGA 356  
  
Qy 1159 GGCTGTCTCTTTGGATGATGCTCTAG 1183  
Db 357 GGCTGTCTCTTTGGATGATGCTCTAG 381

RESULT 14

AL844309  
LOCUS  
DEFINITION  
AL844309 pool\_FLU\_lib\_v\_SPC Homo sapiens cDNA, mRNA sequence.  
VERSION  
AL844309.1 GI:22019091  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE

1 (bases 1 to 608)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,B.J. and Sheridan,E.

TITLE

Homo sapiens EST sequence  
Unpublished (2002)  
Contact: The Sanger Centre  
The Sanger Centre  
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

JOURNAL

Sanger Centre name : scdd10817.1541364  
Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool\_FLU\_lib\_v\_SPC cDNA library. Further information can be found

at <http://www.sanger.ac.uk/Teams/Team69/>.

Location/Qualifiers

source

1..606  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="20"  
/clone\_lib="pool\_FLU\_lib\_v\_SPC"

ORIGIN

Query Match 16.9%; Score 200; DB 9; Length 606;  
Best Local Similarity 99.6%; Pred. No. 8.2e-90;  
Matches 250; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 908 GCCTGTGTGACCATCTACTGAGCTGGCGGATGACATCTGCTGCTACTCAGAGGAGCCC 967  
Db 46 GCCTGTGTGACCATCTACTGAGCTGGCGGATGACATCTGCTGCTACTCAGAGGAGCCC 105  
  
Qy 968 TGTGTCTCTGAGAGGGTGGCCCGCTCCCTGGCAAGGATATACCTTACTGTGACTGTG 1027  
Db 106 TGTGTCTCTGAGAGGGTGGCCCGCTCCCTGGCAAGGATATACCTTACTGTGACTGTG 165  
  
Qy 1028 CAGAGGACACCTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGTGGC 1087  
Db 166 CAGAGGACACCTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTGTTTCTGAAGTGGC 225  
  
Qy 1088 ACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGC 1147  
Db 226 ACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGC 285  
  
Qy 1148 CTGAATGACGA 1158  
Db 286 CTGAATGACGA 296

RESULT 15

AL844308/c  
LOCUS  
DEFINITION  
AL844308 pool\_AK\_lib\_v\_SPD Homo sapiens cDNA, mRNA sequence.  
VERSION  
AL844308.1 GI:22019090  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE

1 (bases 1 to 642)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,B.J. and Sheridan,E.

TITLE

Homo sapiens EST sequence  
Unpublished (2002)  
Contact: The Sanger Centre  
The Sanger Centre  
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

JOURNAL

Sanger Centre name : scdd10816.4004894  
Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool\_AK\_lib\_v\_SPD cDNA library. Further information can be found at <http://www.sanger.ac.uk/Teams/Team69/>.

FEATURES

source

1..642  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="20"  
/clone\_lib="pool\_AK\_lib\_v\_SPD"

ORIGIN

Query Match 9.8%; Score 116; DB 9; Length 642;  
 Best Local Similarity 100.0%; Pred.No. 3.7e-47;  
 Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1068	TCCTGTTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGT	1127
Db	588	TCCTGTTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGGAGT	529
Qy	1128	CCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTTGGATGATGCCTAG	1183
Db	528	CCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTTTTGGATGATGCCTAG	473

Search completed: February 20, 2004, 08:13:51  
 Job time : 2937 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 05:30:56 ; Search time 113 Seconds  
(without alignment)  
5809.796 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agctagagctccaagacc.....tctcttgatgagcctag 1183

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/2/ina/5A-COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B-COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/6A-COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/6B-COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PCITUS-COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	1.7	539	US-09-621-976-10381	Sequence 10381, A
2	20	1.7	675	US-08-707-793A-3	Sequence 3, Appli
3	20	1.7	675	US-08-707-792A-3	Sequence 3, Appli
4	20	1.7	2129	US-09-016-434-1452	Sequence 1452, Ap
5	20	1.7	786431	US-09-751-388-3	Sequence 3, Appli
6	19	1.6	1467	US-09-579-182-2	Sequence 2, Appli
7	19	1.6	1548	US-09-099-053-1	Sequence 1, Appli
8	19	1.6	2771	US-09-016-434-1101	Sequence 1101, Ap
9	18	1.5	1438	US-09-187-331-4	Sequence 4, Appli
10	18	1.5	1438	US-09-470-946-4	Sequence 4, Appli
11	18	1.5	1569	US-09-318-448-8	Sequence 8, Appli
12	18	1.5	3090	US-09-276-531-78	Sequence 78, Appl
13	18	1.5	70000	US-09-551-898-3	Sequence 3, Appli
14	17	1.4	351	US-09-046-479-1	Sequence 1, Appli
15	17	1.4	351	US-08-822-897C-1	Sequence 1, Appli
16	17	1.4	351	US-09-608-810A-3	Sequence 3, Appli
17	17	1.4	351	US-09-404-417A-1	Sequence 1, Appli
18	17	1.4	435	US-09-252-991A-6817	Sequence 6817, Ap
19	17	1.4	439	US-09-222-575-172	Sequence 172, App
20	17	1.4	439	US-09-389-681-172	Sequence 172, App
21	17	1.4	439	US-09-620-405B-172	Sequence 172, App
22	17	1.4	439	US-09-339-338-172	Sequence 172, App
23	17	1.4	439	US-09-433-826B-172	Sequence 172, App
24	17	1.4	439	US-09-604-287A-172	Sequence 172, App
25	17	1.4	439	US-09-285-480-172	Sequence 172, App
26	17	1.4	439	US-09-834-759-172	Sequence 172, App
27	17	1.4	445	US-09-702-705-1598	Sequence 1598, Ap

28 1.4 445 4 US-09-736-457-1598 Sequence 1598, Ap  
29 1.4 445 4 US-09-614-124B-1598 Sequence 1598, Ap  
30 1.4 445 4 US-09-671-325-1598 Sequence 1598, Ap  
31 1.4 541 4 US-09-220-132-10 Sequence 10, Appl  
32 1.4 566 4 US-09-621-976-1574 Sequence 1574, Ap  
33 1.4 627 4 US-09-328-352-1086 Sequence 1086, Ap  
34 1.4 643 4 US-09-833-381-1262 Sequence 1262, Ap  
35 1.4 651 4 US-09-016-434-1255 Sequence 1255, Ap  
36 1.4 674 4 US-09-621-976-87 Sequence 87, Appl  
37 1.4 759 4 US-09-489-039A-5493 Sequence 5493, Ap  
38 1.4 894 4 US-09-434-354-1 Sequence 1, Appl  
39 1.4 941 4 US-09-634-238-89 Sequence 89, Appl  
40 1.4 1089 1 US-08-154-915-1 Sequence 1, Appl  
41 1.4 1089 2 US-08-464-517-37 Sequence 37, Appl  
42 1.4 1089 2 US-08-246-361A-37 Sequence 37, Appl  
43 1.4 1089 3 US-08-463-772-37 Sequence 37, Appl  
44 1.4 1089 5 PCT-US93-09945-1 Sequence 1, Appl  
45 1.4 1747 4 US-09-566-921-66 Sequence 66, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-621-976-10381  
; Sequence 10381, Application US/09621976  
; Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTS and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621.976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 10381

LENGTH: 539

TYPE: DNA

ORGANISM: Homo sapiens

US-09-621-976-10381

Query Match 1.7%; Score 20; DB 4; Length 539;

Best Local Similarity 100.0%; Pred. No. 4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 472 TGTGACCATGGAAGCAGAGA 491

Db 52 TGTGACCATGGAAGCAGAGA 71

##### RESULT 2

US-08-707-793A-3  
; Sequence 3, Application US/08707793A  
; Patent No. 5776696

GENERAL INFORMATION:

APPLICANT: SALOWE, SCOTT P.

TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING

TITLE OF INVENTION: FUSION PROTEINS

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESS: Merck & Co., Inc.

STREET: P.O. Box 2000, 126 E. Lincoln Ave.

CITY: Rahway

STATE: NJ

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/707,793A  
;; FILING DATE: 04-SEP-1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Camara, Valerie J  
;; REGISTRATION NUMBER: 35,090  
;; REFERENCE/DOCKET NUMBER: 19494  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 908-594-3902  
;; TELEFAX: 908-594-4720  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 675 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: Genomic DNA  
US-08-707-793A-3

Query Match 1.7%; Score 20; DB 1; Length 675;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766  
Db 443 CCTTCTCATCCGGGAGGC 462

## RESULT 3

US-08-707-792A-3  
;; Sequence 3, Application US/08707792A  
;; Patent No. 5783399  
;; GENERAL INFORMATION:  
;; APPLICANT: MARCY, ALICE  
;; APPLICANT: SALOME, SCOTT P.  
;; APPLICANT: WISNIEWSKI, DOUGLAS  
;; TITLE OF INVENTION: A HIGH THROUGHPUT ASSAY USING  
;; FUSION PROTEINS  
;; NUMBER OF SEQUENCES: 17  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Merck & Co., Inc.  
;; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
;; CITY: Rahway  
;; STATE: NJ  
;; COUNTRY: USA  
;; ZIP: 07065-0900

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/707,792A  
;; FILING DATE: 04-SEP-1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Camara, Valerie J  
;; REGISTRATION NUMBER: 35,090  
;; REFERENCE/DOCKET NUMBER: 19524  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 908-594-3902  
;; TELEFAX: 908-594-4720  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 675 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: Genomic DNA  
US-08-707-792A-3

Query Match 1.7%; Score 20; DB 1; Length 675;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766  
Db 443 CCTTCTCATCCGGGAGGC 462

## RESULT 4

US-09-016-434-1452  
;; Sequence 1452, Application US/09016434  
;; Patent No. 6500938  
;; GENERAL INFORMATION:  
;; APPLICANT: Janice Au-Young  
;; APPLICANT: Jeffrey J. Seilhamer  
;; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
;; PATHWAY GENE EXPRESSION  
;; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
;; NUMBER OF SEQUENCES: 1490  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
;; STREET: 3174 PORTER DRIVE  
;; CITY: PALO ALTO  
;; STATE: CALIFORNIA  
;; COUNTRY: USA  
;; ZIP: 94304

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/016,434  
;; FILING DATE: HEREWITH  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; CLASSIFICATION:

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Zeller, Karen J.  
;; REGISTRATION NUMBER: 37,071  
;; REFERENCE/DOCKET NUMBER: PA-0002 US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (650) 855-0555  
;; TELEFAX: (650) 845-4166  
;; INFORMATION FOR SEQ ID NO: 1452:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2129 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; LIBRARY: GENBANK  
;; CLONE: 9775207  
US-09-016-434-1452

Query Match 1.7%; Score 20; DB 4; Length 2129;  
Best Local Similarity 100.0%; Pred. No. 4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766  
Db 508 CCTTCTCATCCGGGAGGC 527

```
RESULT 5
US-09-751-389-3
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CLO01067
; CURRENT APPLICATION NUMBER: US/09/751.389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match 1.7%; Score 20; DB 4; Length 786431;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 TGAGCAGGAGGAGGAGCAG 712
DB 412751 TGAGCAGGAGGAGGAGCAG 412770

RESULT 6
US-09-579-182-2
; Sequence 2, Application US/09579182
; Patent No. 6500628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MNI-161
; CURRENT APPLICATION NUMBER: US/09/579.182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-579-182-2

Query Match 1.6%; Score 19; DB 4; Length 1467;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 AGGGGCTTCCTCATCCGG 760
DB 423 AGGGGCTTCCTCATCCGG 441

RESULT 7
US-09-099-053-1
; Sequence 1, Application US/09099053
; Patent No. 6388063
; GENERAL INFORMATION:
; APPLICANT: Greg Plowman
; APPLICANT: Susan Onrust
; APPLICANT: David Markby
; APPLICANT: Sara Courtneidge
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; SAD RELATED DISORDERS
; NUMBER OF SEQUENCES: 28
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099.053
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/049,914
; FILING DATE: June 18, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 235/121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1548 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-099-053-1

Query Match 1.6%; Score 19; DB 4; Length 1548;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 742 AGGGGCTTCCTCATCCGG 760
DB 471 AGGGGCTTCCTCATCCGG 489

RESULT 8
US-09-016-434-1101
; Sequence 1101, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION/DOCKET NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G1256002
US-09-016-434-1101

Query Match 1.6%; Score 19; DB 4; Length 2771;
Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0;

QY 585 CTGAGGATGGAGACTGGTG 603
DB 1305 CTGAGGATGGAGACTGGTG 1323

RESULT 9
US-09-187-331-4
; Sequence 4, Application US/09187331
; Patent No. 6043056
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Gorgone, Gina A.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
; FILE REFERENCE: PF-0631 US
; CURRENT APPLICATION NUMBER: US/09/187,331
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2705267
US-09-187-331-4

QY 1139 TACATCAGCCTGAATGAC 1156
DB 855 TACATCAGCCTGAATGAC 872

RESULT 10
US-09-470-946-4
; Sequence 4, Application US/09470946
; Patent No. 6358923
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Gorgone, Gina A.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSER: INCVTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO

Query Match 1.5%; Score 18; DB 3; Length 1438;
Best Local Similarity 100.0%; Pred. No. 38; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;

QY 1139 TACATCAGCCTGAATGAC 1156
DB 855 TACATCAGCCTGAATGAC 872

RESULT 11
US-09-318-448-8
; Sequence 8, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-8

Query Match 1.5%; Score 18; DB 3; Length 1669;
Best Local Similarity 100.0%; Pred. No. 38; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;

QY 416 AGAAGAAAATCTCTGCCA 433
DB 494 AGAAGAAAATCTCTGCCA 511

RESULT 12
US-09-276-531-78/c
; Sequence 78, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSER: INCVTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
```

STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/276,531  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/079,677  
FILING DATE: March 27, 1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lynn E. Murry, Ph.D.  
REGISTRATION NUMBER: 42,918  
REFERENCE/DOCKET NUMBER: PA-0008 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3090 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAINOT14  
CLONE: 1595762  
US-09-276-531-78

Query Match 1.5%; Score 18; DB 3; Length 3090;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1093 GGAGGAGTCTTCTTCAG 1110  
Db 398 GGAGGAGTCTTCTTCAG 381

RESULT 13  
US-09-851-896-3  
Sequence 3, Application US/09851896  
Patent No. 6410325  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Susan M. Freier  
TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)  
FILE REFERENCE: RTS-0220  
CURRENT FILING DATE: 2001-05-08  
CURRENT APPLICATION NUMBER: US/09/851,896  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 3  
LENGTH: 70000  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-851-896-3

Query Match 1.5%; Score 18; DB 4; Length 70000;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 455 GTCCAGGCCAGGACCT 472  
Db 60708 GTCCAGGCCAGGACCT 60725

RESULT 14  
US-09-046-479-1/c  
Sequence 1, Application: US/09046479  
Patent No. 6291653  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Deisher, Theresa A.  
TITLE OF INVENTION: MOTILIN HOMOLOGS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESS: ZymoGenetics, Inc.  
STREET: 1201 Eastlake Avenue East  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,479  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sawislak, Deborah A.  
REGISTRATION NUMBER: 37,438  
REFERENCE/DOCKET NUMBER: 97-04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-442-6672  
TELEFAX: 206-442-6678  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...351  
OTHER INFORMATION:  
NAME/KEY: sig\_peptide  
LOCATION: 1...69  
OTHER INFORMATION:  
NAME/KEY: mat\_peptide  
LOCATION: 70...351  
OTHER INFORMATION:  
US-09-046-479-1

Query Match 1.4%; Score 17; DB 3; Length 351;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 296 GTCCAGGCCAGCATGC 312  
Db 57 GTCCAGGCCAGCATGC 41

RESULT 15  
US-08-822-897C-1/c  
Sequence 1, Application US/08822897C  
Patent No. 6380158  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
APPLICANT: Deisher, Theresa A.  
TITLE OF INVENTION: MOTILIN HOMOLOGS  
NUMBER OF SEQUENCES: 7

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,897C
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sawislak, Deborah A
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 97-04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...351
OTHER INFORMATION:
NAME/KEY: sig_peptide
LOCATION: 1...69
OTHER INFORMATION:
NAME/KEY: mat_peptide
LOCATION: 70...351
OTHER INFORMATION:
US-08-822-897C-1

```

```

Query Match 1.4%; Score 17; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 296 GTCCAGCCAGAGCATGC 312
Db 57 GTCCAGCCAGAGCATGC 41

```

```

Search completed: February 20, 2004, 08:16:02
Job time : 121 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 07:26:02 ; Search time 486 Seconds  
(without alignments)  
8523.550 Million cell updates/sec

Title: US-09-939-853A-74

Perfect score: 1183

Sequence: 1 agttagctccagagacc.....tctcttgatgacgctag 1183

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2308684 seqs, 1750822206 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW PUB.seq:\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW PUB.seq:\*

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW PUB.seq:\*

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW PUB.seq:\*

8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW PUB.seq:\*

13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*

14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW PUB.seq:\*

17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW PUB.seq:\*

18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	735	62.1	786	14	US-10-043-649-1
2	724	61.2	864	10	US-09-814-353-21302
3	657	55.5	763	9	US-09-867-550-953
4	348	29.4	444	9	US-09-867-550-951
5	341	28.8	875	9	US-09-867-550-1915
6	134	11.3	432	9	US-09-864-761-2829
7	134	11.3	448	9	US-09-864-761-15513
8	96	8.1	96	9	US-09-864-761-19612
9	87	7.4	320	10	US-09-814-353-17314
10	77	6.5	152	10	US-09-814-353-4631
11	77	6.5	152	10	US-09-814-353-10930
12	21	1.8	701	15	US-10-062-674-2188
13	20	1.7	611	15	US-10-027-632-195852
14	20	1.7	672	15	US-10-027-632-107077
15	20	1.7	672	15	US-10-027-632-142058

16	20	1.7	672	15	US-10-027-632-142059	Sequence 142059,
17	20	1.7	672	15	US-10-027-632-142060	Sequence 142060,
18	20	1.7	711	15	US-10-027-632-26286	Sequence 26286, A
19	20	1.7	934	15	US-10-360-238-640	Sequence 640, App
20	20	1.7	2017	15	US-10-062-674-1776	Sequence 1776, Ap
21	20	1.7	2032	15	US-10-366-288-27	Sequence 27, Appl
22	20	1.7	2129	10	US-09-360-706-954	Sequence 954, App
23	20	1.7	2129	15	US-10-305-720-1452	Sequence 1452, Ap
24	20	1.7	786431	14	US-10-412-277-3	Sequence 3, Appl1
25	19	1.6	65	10	US-09-908-975-4510	Sequence 4510, Ap
26	19	1.6	114	9	US-09-864-761-30106	Sequence 30106, A
27	19	1.6	510	14	US-10-072-602B-237	Sequence 237, App
28	19	1.6	599	9	US-09-864-761-13565	Sequence 13565, A
29	19	1.6	1033	9	US-09-789-561-20	Sequence 20, Appl
30	19	1.6	1125	15	US-10-027-632-118578	Sequence 118578,
31	19	1.6	1133	12	US-10-424-599-55347	Sequence 55347, A
32	19	1.6	1467	14	US-10-354-358-11	Sequence 11, Appl
33	19	1.6	1580	10	US-09-976-782-25	Sequence 25, Appl
34	19	1.6	1833	9	US-09-861-846-1	Sequence 1, Appl1
35	19	1.6	2120	15	US-10-094-749-577	Sequence 577, App
36	19	1.6	2771	15	US-10-305-720-1101	Sequence 1101, Ap
37	19	1.6	4207	15	US-10-369-493-46381	Sequence 46381, A
38	19	1.6	4720	15	US-10-062-674-2048	Sequence 2048, Ap
39	18	1.5	204	9	US-09-738-626-3261	Sequence 3261, Ap
40	18	1.5	402	10	US-09-918-995-8609	Sequence 8609, Ap
41	18	1.5	403	9	US-09-983-965-1815	Sequence 1815, Ap
42	18	1.5	474	15	US-10-027-632-195991	Sequence 195991,
43	18	1.5	487	10	US-09-918-995-26739	Sequence 26739, A
44	18	1.5	491	10	US-09-918-995-38019	Sequence 38019, A
45	18	1.5	497	15	US-10-027-632-270409	Sequence 270409,

# ALIGNMENTS

RESULT 1

US-10-043-649-1

Sequence 1, Application US/10043649

Publication No. US2003005924A1

GENERAL INFORMATION:

APPLICANT: Holland, Sacha J.

APPLICANT: Mendenhall, Marcy K.

APPLICANT: Pardo, Jorge

APPLICANT: Spencer, Collin

APPLICANT: Fu, C. Alan

APPLICANT: Luo, Ying

APPLICANT: Pavan, Donald G.

APPLICANT: Mancebo, Helena S.Y.

APPLICANT: Wu, Jun

APPLICANT: Zhou, Xiulan

APPLICANT: Shen, Mary

APPLICANT: Liao, X. Charlene

APPLICANT: Sheng, Ning

TITLE OF INVENTION: Cloning of a No. US2003005924A1el Inhibitor of Antigen-receptor

FILE REFERENCE: A-70219-1/RMS/DHR

CURRENT APPLICATION NUMBER: US/10/043,649

CURRENT FILING DATE: 2002-01-10

PRIOR APPLICATION NUMBER: US 60/260,953

PRIOR FILING DATE: 2001-01-10

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 786

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(786)

OTHER INFORMATION:

US-10-043-649-1

Query Match 62.1%; Score 735; DB 14; Length 786;

```

Best Local Similarity 99.9%; Pred. No. 0;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 398 ATGGAGTCTGCCAGAGAGAAATCTCTGCAAGCCCAAGCTGAGTCTCTGTC 457
Db 1 ATGGAGTCTGCCAGAGAGAAATCTCTGCAAGCCCAAGCTGAGTCTCTGTC 60
QY 458 CAAGCCAGGAGCTGTGACATGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 517
Db 61 CAAGCCAGGAGCTGTGACATGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120
QY 518 GGCAGTTTCCCGAGGTGGCCCGGAGGTGTGCTGAGACTCGGAGAGAGAGAGAGAGAT 577
Db 121 GGCAGTTTCCCGAGGTGGCCCGGAGGTGTGCTGAGACTCGGAGAGAGAGAGAGAT 180
QY 578 ATCGTCTCTGAGATGGAGACTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAT 637
Db 181 ATCGTCTCTGAGATGGAGACTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
QY 638 AACATCCCGAGCGTCCAGTGGGCAAGTCTCCCATGGTGGTGTATGAGGGCTGAGC 697
Db 241 AACATCCCGAGCGTCCAGTGGGCAAGTCTCCCATGGTGGTGTATGAGGGCTGAGC 300
QY 698 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 757
Db 301 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 360
QY 758 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 817
Db 361 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 420
QY 818 TCCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 877
Db 421 TCCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
QY 878 TCACGCGGCTCACCTTCCCTCTACTCCAGGAGAGAGAGAGAGAGAGAGAGAT 937
Db 481 TCACGCGGCTCACCTTCCCTCTACTCCAGGAGAGAGAGAGAGAGAGAGAGAT 540
QY 938 GATGACATCTGCTGCTACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 997
Db 541 GATGACATCTGCTGCTACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 600
QY 998 GGCAGAGATATACCCCTACTCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAT 1057
Db 601 GGCAGAGATATACCCCTACTCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAT 660
QY 1058 GACAGTCCCTCTCTGTTTGAAGCTGCCAGAGAGAGAGAGAGAGAGAGAT 1117
Db 661 GACAGTCCCTCTCTGTTTGAAGCTGCCAGAGAGAGAGAGAGAGAGAGAT 720
QY 1118 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGAGAGAGAGAGAGAT 1177
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGATGAGAGAGAGAGAGAT 780
QY 1178 GCCTAG 1183
Db 781 GCCTAG 786

```

RESULT 2

```

US-09-814-353-21302
; Sequence 21302, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353

```

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21302
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 32, 862, 863, 864
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21302

Query Match 61.2%; Score 724; DB 10; Length 864;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 774; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GCTAGAGCTCCAGGAGCCACGCGCTGTCTGTGACAGAGCTCAAGAGGCCCTGGGC 61
Db 54 GCTAGAGCTCCAGGAGCCACGCGCTGTCTGTGACAGAGCTCAAGAGGCCCTGGGC 113
QY 62 CTTCCCTCCCTGGCTGGTGTGTTGGAGGGTTCCTCCAGTCCAGATCTCTAAGAGC 121
Db 114 CTTCCCTCCCTGGCTGGTGTGTTGGAGGGTTCCTCCAGTCCAGATCTCTAAGAGC 173
QY 122 ATGGGACAGCTGATCATCTCCCTGGTGTACAACTGCTGACTGCAGACAGATCTGAGCTA 181
Db 174 ATGGGACAGCTGATCATCTCCCTGGTGTACAACTGCTGACTGCAGACAGATCTGAGCTA 233
QY 182 CCCAAACCAACACCTAGCCTCTCCCTGAGATCTCCCTCCAGGCTGAGAGAGTCTCGGGT 241
Db 234 CCCAAACCAACACCTAGCCTCTCCCTGAGATCTCCCTCCAGGCTGAGAGAGTCTCGGGT 293
QY 242 CCTAGACCAAGAGACTGGGAGACTTCCAGAGGGCCCCCAAGCCCTTAACCTGTCAG 301
Db 294 CCTAGACCAAGAGACTGGGAGACTTCCAGAGGGCCCCCAAGCCCTTAACCTGTCAG 353
QY 302 CCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGTATGACAAACCAATTTCCC 361
Db 354 CCAGAGCATGCGTCTCAGCAGAGCTGTCTTCCCAAGCCCTTGTATGACAAACCAATTTCCC 413
QY 362 TCGATGATGTGTTCTGAGTGTCTGAGAGAACAATGGGAGTCTGCCAGAGAGAGA 421
Db 414 TCGATGATGTGTTCTGAGTGTCTGAGAGAACAATGGGAGTCTGCCAGAGAGAGA 473
QY 422 AAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTGTCTCCAGGCGCAGGAGCTGTGACCATG 481
Db 474 AAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTGTCTCCAGGCGCAGGAGCTGTGACCATG 533
QY 482 GAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 541
Db 534 GAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 593
QY 542 GCGAGCTGTGCTGAGACTCGGGAGAGCATTGACATCGTCTCTGAGGATGAGAGACTGG 601
Db 594 GCGAGCTGTGCTGAGACTCGGGAGAGCATTGACATCGTCTCTGAGGATGAGAGACTGG 653
QY 602 TGAACCGTGTCTGTAAGTCTCAGCAGAGAGATTAACATCCCGAGCGTCCACGTGGGC 661
Db 654 TGAACCGTGTCTGTAAGTCTCAGCAGAGAGATTAACATCCCGAGCGTCCACGTGGGC 713

```

QY 662 AAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGAGAAAGCAGAGGAAGTGTCTG 721  
Db 714 AAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGAGAAAGCAGAGGAAGTGTCTG 773  
QY 722 TTGTTACCTGGGAACCCCTGGAGGGGCTTCTCATCCGGAGAGCCAGACCAAGGA 776  
Db 774 TTGTTACCTGGGAACCCCTGGAGGGGCTTCTCATCCGGAGAGCCAGACCAAGGA 828

## RESULT 3

US-09-867-550-953  
; Sequence 953, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; PRIOR FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 953  
; LENGTH: 763  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-867-550-953

Query Match 55.5%; Score 657; DB 9; Length 763;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 117 GGAGCATGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGAGACAGATGCTG 176  
Db 5 GGAGCATGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACTGAGACAGATGCTG 64  
QY 177 AGCTACCCAAACCAACACCTGACCTCTCCCTGAAGATCTCCAGGCTGAGAGGTTCTG 236  
Db 65 AGCTACCCAAACCAACACCTGACCTCTCCCTGAAGATCTCCAGGCTGAGAGGTTCTG 124  
QY 237 GGTGTCCTAGGACCAAGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACTG 296  
Db 125 GATGTCCTAGGACCAAGACACTGGCAGACTTCCAGAGGGGCCCAAGCCCTAACTG 184  
QY 297 TCCAGCCAGAGCATGGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAAT 356  
Db 185 TCCAGCCAGAGCATGGTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAAT 244  
QY 357 TTCCCTCGATGATGCTTCTGATGCTCTGCTGAGAACCAATGGAAAGTGTGCCAGCA 416  
Db 245 TTCCCTCGATGATGCTTCTGATGCTCTGCTGAGAACCAATGGAAAGTGTGCCAGCA 304  
QY 417 GAAGAAATCTCTGCCAAGCCCAAGCTTCTGCTGAGAACCAATGGAAAGTGTGCCAGCA 476  
Db 305 GAAGAAATCTCTGCCAAGCCCAAGCTTCTGCTGAGAACCAATGGAAAGTGTGCCAGCA 364  
QY 477 CCATGGAAGCAGAGAGAGCAAGCCACAGCCCTGGCCCTGGGCAAGTTTCCCGCAGGTG 536  
Db 365 CCATGGAAGCAGAGAGAGCAAGCCACAGCCCTGGCCCTGGGCAAGTTTCCCGCAGGTG 424  
QY 537 GCCCGCCGAGCTGTCGTGAGCTCGGGAGCCATTGACCATGCTCTCTGAGATGGAG 596  
Db 425 GCCCGCCGAGCTGTCGTGAGCTCGGGAGCCATTGACCATGCTCTCTGAGATGGAG 484  
QY 597 ACTGTTGAGACGGTGTCTGTAAGTCTCAGGACAGAGTATAACATCCCAAGCCTCCACG 656  
Db 485 ACTGTTGAGACGGTGTCTGTAAGTCTCAGGACAGAGTATAACATCCCAAGCCTCCACG 544

QY 657 TGGCAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGAGAAAGCAGAGGAAC 716  
Db 545 TGGCAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGCAGGAGAGAAAGCAGAGGAAC 604  
QY 717 TGCTGTTGTTACCTGGGAACCCCTGGAGGGGCTTCTCATCCGGAGAGCCAGACCAAGGA 776  
Db 605 TGCTGTTGTTACCTGGGAACCCCTGGAGGGGCTTCTCATCCGGAGAGCCAGACCAAGGA 664  
QY 777 GAGGCTCTTACTCTCTCTGATCGGCTCAGCGCCCTGTCATCTCTGGACCGGATCAGAC 836  
Db 665 GAGGCTCTTACTCTCTCTGATCGGCTCAGCGCCCTGTCATCTCTGGACCGGATCAGAC 724  
QY 837 ACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACA 875  
Db 725 ACTACAGGATCCACTGCTTGAACAATGGCTGGCTGTACA 763

## RESULT 4

US-09-867-550-951  
; Sequence 951, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; PRIOR FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 951  
; LENGTH: 444  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-867-550-951

Query Match 29.4%; Score 348; DB 9; Length 444;

Best Local Similarity 100.0%; Pred. No. 7.5e-167;

Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 CCTAGCACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTAACTGTCCAG 301  
Db 1 CCTAGCACCAAGGACACTGGCAGACTTCCAGAGGGGCCCCCAAGCCCTAACTGTCCAG 60  
QY 302 CCAGAGATCGGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCC 361  
Db 61 CCAGAGATCGGCTCTCAGCAGAGCTGTCTTCCCAAGCCTTTGATGACAAACCAATTTCCC 120  
QY 362 TCATGATGTCTTCTGAGTGTCTCTCAGAGAAACAAATGGAGTCTGCCAGCAGCAAGA 421  
Db 121 TCATGATGTCTTCTGAGTGTCTCTCTGAGAACAAATGGAGTCTGCCAGCAGCAAGA 180  
QY 422 AATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCCAAAGCCAGGACCTGTGACCATG 481  
Db 181 AATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTCCAAAGCCAGGACCTGTGACCATG 240  
QY 482 GAAGCAGAGAGAGCAAGCCACAGCCGTGGCCCTGGGCAAGTTTCCCGCAGGTGCCCCG 541  
Db 241 GAAGCAGAGAGAGCAAGCCACAGCCGTGGCCCTGGGCAAGTTTCCCGCAGGTGCCCCG 300  
QY 542 GCGAGCTCTCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAG 589  
Db 301 GCGAGCTCTCTGAGACTCGGGAGCCATTGACCATGCTCTCTGAG 348

## RESULT 5

```

US-09-867-550-1915
; Sequence 1915, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; FILE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1915
; LENGTH: 875
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: Wherein n is one of a or t or c or g
US-09-867-550-1915

Query Match      28.8%; Score 341; DB 9; Length 875;
Best Local Similarity 100.0%; Pred. No. 2.5e-163; Indels 0; Gaps 0;
Matches 341; Conservative 0; Mismatches 0;

QY 843 GGATCCACTGCTTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCTCAC 902
DB 2 GGATCCACTGCTTGACAAATGGCTGGCTGTACATCTCACCGCGCTCACCTTCCCTCTCAC 61
QY 903 TCCAGGCCCTGGTGGACCAATTAATCTGAGCTGGGGAGTACATCTGCTGCTACTCAAG 962
DB 62 TCCAGGCCCTGGTGGACCAATTAATCTGAGCTGGGGAGTACATCTGCTGCTACTCAAG 121
QY 963 AGCCCTGTCTCTGTCAGAGGCTGCGCGCTCCCTGGCAAGATATACCCCTACTCTGA 1022
DB 122 AGCCCTGTCTCTGTCAGAGGCTGCGCGCTCCCTGGCAAGATATACCCCTACTCTGA 181
QY 1023 CTGTCAGAGACACCACTCAACTGGAAGAGCTGGAGCTGCTCTCTGTTTCTGAAG 1082
DB 182 CTGTCAGAGACACCACTCAACTGGAAGAGCTGGAGCTGCTCTCTGTTTCTGAAG 241
QY 1083 CTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGTCTCCGGAGTCCCTCAGCTTCTACA 1142
DB 242 CTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGTCTCCGGAGTCCCTCAGCTTCTACA 301
QY 1143 TCAGCCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 1183
DB 302 TCAGCCTGAATGACAGGCTGTCTCTTTGGATGATGCTAG 342

RESULT 6
US-09-864-761-2829
; Sequence 2829, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

```

```

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Arcomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2829
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL031662.24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ST474, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
US-09-864-761-2829

Query Match      11.3%; Score 134; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 8.3e-58;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 929 GAGCTGGCGGATGACATCTGCTGCTTCAAGAGCCCTGTGCTCTGCAGAGGCTGCG 988
DB 270 GAGCTGGCGGATGACATCTGCTGCTTCAAGAGCCCTGTGCTCTGCAGAGGCTGCG 329
QY 989 CGCTCCCTGGCAAGGATATACCCCTACTGTGCTGTCAGAGGACACCCTCAACTGG 1048
DB 330 CGCTCCCTGGCAAGGATATACCCCTACTGTGCTGTCAGAGGACACCCTCAACTGG 389
QY 1049 AAAGAGCTGGACAG 1062
DB 390 AAAGAGCTGGACAG 403

RESULT 7
US-09-864-761-15513

```



```
; OTHER INFORMATION: EST_HUMAN HIT: A1125308.1, EVALUATE 2.10e-01
US-09-864-761-19612

Query Match
Best Local Similarity 8.1%; Score 96; DB 9; Length 96;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 967 CTGTCTCTCGAGAGGCTGGCCCTCCCTGGCAAGATATACCCCTACCTGTGACTGT 1026
DB 1 CTGTCTCTCGAGAGGCTGGCCCTCCCTGGCAAGATATACCCCTACCTGTGACTGT 60

QY 1027 GCAGAGGACCACTCACTCACTGGAAGAGCTGCACAG 1062
DB 61 GCAGAGGACCACTCACTCACTGGAAGAGCTGCACAG 96

RESULT 9
US-09-814-353-17314
; Sequence 17314, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17314
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-17314

Query Match
Best Local Similarity 7.4%; Score 87; DB 10; Length 320;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 TGTGCTGAGACTCGGGAGCCATTGACCATCTCTGAGGATGAGACTGGTGACGG 608
DB 103 TGTGCTGAGACTCGGGAGCCATTGACCATCTCTGAGGATGAGACTGGTGACGG 162

QY 609 TGTGCTGAGACTCGGGAGCCATTGACCATCTCTGAGGATGAGACTGGTGACGG 635
DB 163 TGTGCTGAGACTCGGGAGCCATTGACCATCTCTGAGGATGAGACTGGTGACGG 189

RESULT 10
US-09-814-353-4631
; Sequence 4631, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
```

```
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4631
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 17, 102, 112
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-4631

Query Match
Best Local Similarity 6.5%; Score 77; DB 10; Length 152;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 TGTGCTGAGACTCGGGAGCCATTGACCATCTCTGAGGATGAGACTGGTGACGG 608
DB 25 TGTGCTGAGACTCGGGAGCCATTGACCATCTCTGAGGATGAGACTGGTGACGG 84

QY 609 TGTGCTGAGACTCGGGAGCCATTGACCATCTCTGAGGATGAGACTGGTGACGG 625
DB 85 TGTGCTGAGACTCGGGAGCCATTGACCATCTCTGAGGATGAGACTGGTGACGG 101

RESULT 11
US-09-814-353-10930
; Sequence 10930, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10930
; LENGTH: 152
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
;
; NAME/KEY: misc feature
; LOCATION: 17, 102, 112
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-10930

Query Match
Best Local Similarity 6.5%; Score 77; DB 10; Length 152;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 TGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGAGACTGCTGTGACGG 608
Db 25 TGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGAGACTGCTGTGACGG 84

QY 609 TGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGAGACTGCTGTGACGG 625
Db 85 TGTGCTGAGACTCGGGAGCCATTGACCATCGTCTCTGAGGATGAGACTGCTGTGACGG 101

RESULT 12
US-10-062-674-2188/c
; Sequence 2188, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 2188
; LENGTH: 701
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040005559A1 893157.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (701)
; OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-2188

Query Match
Best Local Similarity 1.8%; Score 21; DB 15; Length 701;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 TGGCCCTGGGCGAGTTTCCCGG 530
Db 285 TGGCCCTGGGCGAGTTTCCCGG 265

RESULT 13
US-10-027-632-195852/c
; Sequence 195852, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107077
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-195852

Query Match
Best Local Similarity 1.7%; Score 20; DB 15; Length 672;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766
Db 71 CCTTCTCATCCGGGAGGC 90

RESULT 15
US-10-027-632-142058
; Sequence 142058, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
;
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195852
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-195852

Query Match
Best Local Similarity 1.7%; Score 20; DB 15; Length 611;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 CACTGGCAGACTTCCAGAAG 275
Db 503 CACTGGCAGACTTCCAGAAG 484

RESULT 14
US-10-027-632-107077
; Sequence 107077, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107077
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-107077

Query Match
Best Local Similarity 1.7%; Score 20; DB 15; Length 672;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CCTTCTCATCCGGGAGGC 766
Db 71 CCTTCTCATCCGGGAGGC 90

RESULT 15
US-10-027-632-142058
; Sequence 142058, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 142058

; LENGTH: 672

; TYPE: DNA

; ORGANISM: Human

; US-10-027-632-142058

Query Match 1.7%; Score 20; DB 15; Length 672;

Best Local Similarity 100.0%; Pred. No. 9.5;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 747 CTTCTCATCCGGGAGAGC 766

Db 71 CTTCTCATCCGGGAGAGC 90

Search completed: February 20, 2004, 09:27:51

Job time : 487 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 08:14:02 ; Search time 93 Seconds

(without alignments)

792.956 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLPRLRRKSLRPSLSSV.....RESLSFVSLNDAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1353	100.0	261	5	Aau91308 Human pro
2	1347	99.6	261	5	Aao15457 Human mod
3	1347	99.6	261	5	Abp52187 Human Src
4	1347	99.6	261	5	Aae26357 Human Src
5	1269	93.8	248	3	Aab42993 Human ORF
6	1028	76.0	259	5	Aao15456 Mouse mod
7	934.5	69.1	210	5	Aao15458 Mouse mod
8	923.5	68.3	210	5	Abp64788 Human pro
9	826	61.0	159	5	Abp64107 Human ORF
10	586	43.3	113	5	Abp64588 Human ORF
11	481.5	35.6	276	5	Abp52188 Human Src
12	481.5	35.6	276	6	Abu79103 Src-like
13	481.5	35.6	276	6	Abp59698 Human Src
14	470.5	34.8	315	4	Aau31072 Novel hum
15	452.5	33.4	96	5	Abp64106 Human ORF
16	370.5	27.4	512	7	Add19014 Human dis
17	360.5	26.6	505	4	Aab99332 Human tyr
18	360.5	26.6	505	7	Abw01407 Human hae
19	348	25.7	508	3	Aay49420 PKA subst
20	348	25.7	509	6	Abp58699 Human can
21	348	25.7	509	6	Ade40449 Human pro
22	340	25.1	363	6	Abp59690 Human p56
23	340	25.1	437	5	Abg79672 Tumour in
24	340	25.1	508	3	Aab37700 Human lym
25	340	25.1	508	7	Ade58802 Human pro

26	340	25.1	508	7	Ade58799 Human pro
27	340	25.1	509	7	Abp56202 Human lym
28	338	25.0	70	4	Abg05994 Novel hum
29	337	24.9	567	5	Abg79673 Tumour in
30	332	24.5	517	4	Abp57957 Drosophil
31	317	23.4	541	5	Aau74614 Perinucle
32	315.5	23.3	536	7	Add45918 Human pro
33	315.5	23.3	537	7	Add14129 Human src
34	315.5	23.3	543	4	Aab84663 Amino aci
35	315.5	23.3	543	4	Abg10302 Novel hum
36	315.5	23.3	543	6	Ada00843 Human src
37	313.5	23.2	496	2	Aay29658 Human src
38	313.5	23.2	496	4	Aau08734 Xenopus l
39	313.5	23.2	496	4	Aau08730 Xenopus l
40	313.5	23.2	496	4	Aau08735 Xenopus l
41	313.5	23.2	543	2	Aay24421 Human yes
42	312.5	23.1	529	7	Ade31767 Human 145
43	306.5	22.7	551	4	Abg22264 Novel hum
44	303	22.4	458	7	Ade99048 Human kpp
45	286.5	21.2	533	3	Aay44449 Mutant ch

## ALIGNMENTS

### RESULT 1

AAU91308

ID AAU91308 standard; protein; 261 AA.

XX AC AAU91308;

XX DT 18-JUN-2002 (first entry)

XX DE Human protein NOV13.

XX KW Human; NOVX; gene therapy; cardiomyopathy; atherosclerosis;

XX KW cell signal processing disorder; metabolic pathway modulation disorder;

XX KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;

XX KW uterus cancer; immune response; graft-versus-host disease;

XX KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;

XX KW hypertension; congenital heart defects; multiple sclerosis; inflammation;

XX KW Albright hereditary osteodysplasia.

XX OS Homo sapiens.

XX PN WO200216599-A2.

XX PD 28-FEB-2002.

XX PF 27-AUG-2001; 2001WO-US026510.

XX PR 25-AUG-2000; 2000US-0228191P.

XX PR 08-FEB-2001; 2001US-0267300P.

XX PR 20-FEB-2001; 2001US-0269961P.

XX PR 20-MAR-2001; 2001US-0277337P.

XX PA (CURA-) CURAGEN CORP.

XX PA (CORT-) COR THERAPEUTICS INC.

XX PI Burgess CB, Conley PB, Grosse WM, Kekuda R, Shinkets RA;

XX PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;

XX DR MPI; 2002-280937/32.

XX DR N-PSDB; ABK61465.

XX PT New polypeptides for treating or preventing a disorder associated with

XX PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.

XX PS Claim 3; Page 98; 263pp; English.

XX CC The invention relates to an isolated polypeptide (NOVX) a mature form of

XX CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide

XX CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,

15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it and antibody against it, are useful for treating or preventing (e.g. by gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy, atherosclerosis, a disorder related to cell signal processing and metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide and nucleic acids are also useful for determining the presence of a predisposition to the disease. The NOVX nucleic acid and polypeptide are especially useful in therapeutic or prophylactic applications for disorders associated with aberrant NOVX expression or activity, e.g. cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus cancer), immune response, graft-versus-host disease, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension, congenital heart defects, multiple sclerosis, inflammation or Albritght hereditary osteodysrophy and many other diseases listed in the specification. The DNA encoding the protein is useful in gene therapy for treating the conditions. This is also useful in detection assays, chromosome mapping, tissue typing, diagnostic or prognostic assays, or for developing a powerful assay system for functional analysis of various human disorders, as well as in diagnostic applications. The present sequence represents a NOVX protein

Sequence 261 AA;

Query Match 100.0%; Score 1353; DB 5; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-130;  
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSLSRKRKSLPSPSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 DB 1 MGSLSRKRKSLPSPSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60

QY 61 IVSEGDGWTVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKAEELLLPGNPGGAFLI 120  
 DB 61 IVSEGDGWTVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKAEELLLPGNPGGAFLI 120

QY 121 RESQTRRGSYLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFPSLQALVDHYSELA 180  
 DB 121 RESQTRRGSYLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFPSLQALVDHYSELA 180

QY 181 DDICLLKEPCVLRAGPLPGKDIPLVTVQRTPLNWKELDSSLLFSEAATGEESLLSEG 240  
 DB 181 DDICLLKEPCVLRAGPLPGKDIPLVTVQRTPLNWKELDSSLLFSEAATGEESLLSEG 240

QY 241 LRESLSFYISLNDEAVSLDDA 261  
 DB 241 LRESLSFYISLNDEAVSLDDA 261

# RESULT 2

AAO15457  
 ID AAO15457 standard; protein; 261 AA.

AC AAO15457;

DT 03-OCT-2002 (first entry)

DE Human modulator of antigen receptor signalling (MARS) protein.

KW Human; gene therapy; modulator of antigen receptor signalling; MARS;  
 KW tumour suppressor gene; Src-like adaptor protein; SLAP;  
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.

OS Homo sapiens.

XX WO200242452-A2.

PN 30-MAY-2002.

XX 26-NOV-2001; 2001WO-CA001662.

XX 27-NOV-2000; 2000CA-02324663.

XX

PA (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX McGlade JC, Loreto MP;

XX WPI; 2002-566564/60.

XX N-PSDB; AAL44089.

XX New isolated modulator of antigen receptor signaling protein or its fragment, useful for treating malignant disorders such as myeloid malignancies, autoimmune disorders and myeloproliferative disorders.

XX Claim 7; Fig 9A; 110pp; English.

XX The invention comprises the amino acid and coding sequences of modulator of antigen receptor signalling (MARS) proteins. The MARS protein is a putative tumour suppressor gene and exhibits structural and sequence similarity to the Src-like adaptor protein (SLAP). The MARS DNA and protein sequences of the invention are useful for the treatment of myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune disorders, immunosuppression, myeloproliferative disorders and malignancies related to the de-regulation of tyrosine kinases (e.g. breast cancer). The present amino acid sequence represents a human MARS protein

XX Sequence 261 AA;

Query Match 99.6%; Score 1347; DB 5; Length 261;  
 Best Local Similarity 99.6%; Pred. No. 6.8e-130;  
 Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSLSRKRKSLPSPSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 DB 1 MGSLSRKRKSLPSPSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60

QY 61 IVSEGDGWTVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKAEELLLPCNPGGAFLI 120  
 DB 61 IVSEGDGWTVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKAEELLLPCNPGGAFLI 120

QY 121 RESQTRRGSYLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFPSLQALVDHYSELA 180  
 DB 121 RESQTRRGSYLSVRLSRPASWDRIHYRIHCLDNGWLYISPRITFPSLQALVDHYSELA 180

QY 181 DDICLLKEPCVLRAGPLPGKDIPLVTVQRTPLNWKELDSSLLFSEAATGEESLLSEG 240  
 DB 181 DDICLLKEPCVLRAGPLPGKDIPLVTVQRTPLNWKELDSSLLFSEAATGEESLLSEG 240

QY 241 LRESLSFYISLNDEAVSLDDA 261  
 DB 241 LRESLSFYISLNDEAVSLDDA 261

# RESULT 3

ABP52187

ID ABP52187 standard; protein; 261 AA.

XX AC ABP52187;

XX 15-OCT-2002 (first entry)

XX Human Src-like inhibitory molecule (SLIM) protein.

XX Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;  
 KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;  
 KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;  
 KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
 KW chronic inflammatory disorder; autoimmune disorder; transplant rejection.

XX Homo sapiens.

XX WO200255707-A2.

XX 18-JUL-2002.

XX

PF 10-JAN-2002; 2002WO-US000718.  
 XX  
 PR 10-JAN-2001; 2001US-0260953P.  
 XX  
 PA (RIGE-) RIGEL PHARM INC.  
 XX  
 PI Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
 XX  
 DR WPI; 2002-575432/61.  
 DR N-PSDB; ABQ/4343.  
 XX  
 PT New src-like inhibitory molecule protein, useful for treating  
 PT immunodeficiency disorders and inflammatory disorders, comprises N-  
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.  
 XX  
 PS Claim 7; Fig 2A; 9lpp; English.  
 XX  
 CC The present sequence represents the human Src-like inhibitory molecule  
 CC (SLIM) protein (I). The present invention describes a SLIM protein  
 CC comprising an N-terminal myristylation sequence, an N-terminal SH2  
 CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising  
 CC an N-terminal myristylation sequence and an N-terminal SH2 domain which  
 CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and  
 CC anti-Hiv activities, and can be used as a modulator of lymphocyte  
 CC activation, and of ubiquitination of a Cbl target protein, and in gene  
 CC therapy. (I) is useful for screening a bioactive agent capable of binding  
 CC to SLIM. (I) is also useful for screening a bioactive agent capable of  
 CC modulating SLIM binding. (I) or its fragments is useful in the study or  
 CC in the treatment of conditions which involves this function or  
 CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or  
 CC prevent SLIM associated disorders. (I) or the polynucleotide encoding it  
 CC (II) is useful for modulating leukocyte and/or platelet activation, for  
 CC modulating antigen receptor-induced signalling and activation in  
 CC leukocyte and/or platelet and for modulating antigen receptor-induced  
 CC signalling and activation in lymphocytes and/or mast cells. (I) or (II)  
 CC is also useful for modulating the basal activity of lymphocytes. (I) or  
 CC (II) is useful in the treatment of immunodeficiency disorders, such as  
 CC acquired immunodeficiency syndrome (AIDS), for the prevention and  
 CC treatment of acute inflammatory disorders, chronic inflammatory  
 CC disorders, autoimmune disorder and transplant rejection  
 XX  
 SQ Sequence 261 AA;  
 Query Match 99.6%; Score 1347; DB 5; Length 261;  
 Best Local Similarity 99.6%; Pred. No. 6.8e-130;  
 Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGSLPSRRKSLPSLSVSSVQGGPVTWEAERSKATAVALGSPAGGPAELSLRLGEPLT 60  
 Db 1 MGSLPSRRKSLPSLSVSSVQGGPVTWEAERSKATAVALGSPAGGPAELSLRLGEPLT 60  
 QY 61 IVSEGDWMTVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKABELLLLPNGGAPLI 120  
 Db 61 IVSEGDWMTVLSEVSGREYNIPSVHVAKVSHGWLVEGLSREKABELLLLPNGGAPLI 120  
 QY 121 RRSQTRRGYSLSVRLSPASWDRIHVRHICLDNGWLYISPRITFPSLQALVPHYSELA 180  
 Db 121 RRSQTRRGYSLSVRLSPASWDRIHVRHICLDNGWLYISPRITFPSLQALVPHYSELA 180  
 QY 181 DPICLLKEPCVLQAGLPFGKDIPVTVTQRTPLNWKELDSLLFSAATGESLLSEG 240  
 Db 181 DPICLLKEPCVLQAGLPFGKDIPVTVTQRTPLNWKELDSLLFSAATGESLLSEG 240  
 QY 241 LRESLSFYLSINDEAVSLDDA 261  
 Db 241 LRESLSFYLSINDEAVSLDDA 261  
 RESULT 4  
 ID AAE26357  
 XX AAE26357 standard; protein; 261 AA.

AC AAE26357;  
 XX 13-DEC-2002 (first entry)  
 DE Human Src-Like Adapter Protein-2 (hSLAP-2).  
 XX  
 KW Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 35..90  
 FT /note= "SH3 domain"  
 FT Modified-site 80  
 FT /note= "Tyrosine phosphorylation site"  
 FT Domain 94..176  
 FT /note= "SH2 domain"  
 FT Modified-site 96  
 FT /note= "Tyrosine phosphorylation site"  
 FT Modified-site 130  
 FT /note= "Tyrosine phosphorylation site"  
 FT Modified-site 148  
 FT /note= "Tyrosine phosphorylation site"  
 FT Modified-site 159  
 FT /note= "Tyrosine phosphorylation site"  
 FT Modified-site 176  
 FT /note= "Tyrosine phosphorylation site"  
 PN WO200242457-A1.  
 XX 30-MAY-2002.  
 XX 20-NOV-2001; 2001WO-US043367.  
 XX 22-NOV-2000; 2000US-0252545P.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;  
 DR WPI; 2002-463632/49.  
 DR N-PSDB; AAD43980.  
 XX Novel substantially purified human SH2/SH3-domain-containing adapter  
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic  
 PT intervention in immunological and inflammatory disorders and cancer.  
 XX  
 PS Claim 9; Fig 2; 85pp; English.  
 XX  
 CC The invention relates to a substantially purified human SH2/SH3-domain-  
 CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-  
 CC 2). The invention is useful for treating an immune disorder involving  
 CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is  
 CC useful for screening for antagonists or inhibitors of the interaction of  
 CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or  
 CC preventing diseases or disorders associated with aberrant or uncontrolled  
 CC cellular signal transduction, for determining those cellular signalling  
 CC molecules which associate with hSLAP-2 and which provide critical signals  
 CC for cell activation, and as effectors in methods to affect T- cell  
 CC activation. The invention is useful in screening assays to identify and  
 CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for  
 CC potential use to treat autoimmune diseases which may be caused by  
 CC hyperactivated B cells, as well as to treat diseases which may be caused  
 CC by hyperactivated T cells, in addition to other immune system related  
 CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,  
 CC inflammation disorders, diseases and conditions, rheumatoid arthritis,  
 CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's

CC and ulcerative colitis), allergies, particularly those involving  
 CC hyperactivity of B-cells and T-cells, or other immune cells, such as  
 CC mast cells or eosinophils, autoimmune diseases such as systemic lupus  
 CC erythematosus and multiple sclerosis, pulmonary diseases including  
 CC asthma, acute respiratory distress syndrome, and chronic obstructive  
 CC pulmonary disorder, tissue/organ rejection and cancer. The invention is  
 CC useful in gene therapy. The present sequence is human SLAP-2 protein  
 XX  
 XX Sequence 261 AA;

Query Match 99.6%; Score 1347; DB 5; Length 261;  
 Best Local Similarity 99.6%; Pred. No. 6.8e-130;  
 Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGSLSRRKSLSPSLSSSVQGGPVTMEAEKSKATAVAIGSPAGGPAELSLRLGPLELT 60  
 Db 1 MGSLSRRKSLSPSLSSSVQGGPVTMEAEKSKATAVAIGSPAGGPAELSLRLGPLELT 60  
 QY 61 IVSEGDWMTVLSEVSGREYNTPSVHVGVKSHGWLVEGLSREKABELLLPFGGAPLI 120  
 Db 61 IVSEGDWMTVLSEVSGREYNTPSVHVGVKSHGWLVEGLSREKABELLLPFGGAPLI 120  
 QY 121 RESQTRRGYSLSVRLSPASWDRIHRYHICLDNGWLYISPRLTFFPSLQALVDHYSCLA 180  
 Db 121 RESQTRRGYSLSVRLSPASWDRIHRYHICLDNGWLYISPRLTFFPSLQALVDHYSCLA 180  
 QY 181 DDICLLKEPCVLRAGPLPGKDIPLPVTQRTPLNWKELSSLLFSEAAATGEESLLSEG 240  
 Db 181 DDICLLKEPCVLRAGPLPGKDIPLPVTQRTPLNWKELSSLLFSEAAATGEESLLSEG 240  
 QY 241 LRESLSFYISLNDVAVSLDDA 261  
 Db 241 LRESLSFYISLNDVAVSLDDA 261

RESULT 5  
 AAB42993  
 ID AAB42993 standard; protein; 248 AA.  
 AC AAB42993;  
 XX  
 XX 08-FEB-2001 (first entry)  
 XX  
 XX Human ORFX ORF2757 polypeptide sequence SEQ ID NO:5514.  
 XX  
 XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnery; antipariatic; antiparkinsonian; nontropic; neuroprotective;  
 KW immunostimulant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW hypotensive; dermatological; immunosuppressive; antidiabetic;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antiinflammatory;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX  
 XX Homo sapiens.

AAO15456  
 ID AAO15456 standard; protein; 259 AA.  
 XX  
 XX AAO15456;  
 AC  
 XX  
 XX 03-OCT-2002 (first entry)  
 DT  
 XX  
 XX Mouse modulator of antigen receptor signalling (MARS) protein.  
 DE  
 XX  
 XX Mouse; gene therapy; modulator of antigen receptor signalling; MARS;  
 KW  
 KW tumour suppressor gene; Scr-like adaptor protein; SLAP;

PA (CURA-) CURAGEN CORP.  
 XX  
 XX Shimkets RA, Leach M;  
 XX  
 XX WPI; 2000-602362/57.  
 DR N-PSDB; AAC77202.  
 DR  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 XX Claim 11; Page 4693-4694; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 CC antipariatic; antiparkinsonian; nontropic; neuroprotective; osteopathic;  
 CC immunostimulant; antidiabetic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antirheumatic; antiinflammatory;  
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 XX Sequence 248 AA;

Query Match 93.8%; Score 1269; DB 3; Length 248;  
 Best Local Similarity 98.8%; Pred. No. 6.8e-122;  
 Matches 244; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 15 SLSSSVQGGPVTMEAEKSKATAVAIGSPAGGPAELSLRLGPLELTIVSEGDWMTVLSE 74  
 Db 2 SLSSSVQGGPVTMEAEKSKATAVAIGSPAGGPAELSLRLGPLELTIVSEGDWMTVLSE 61  
 QY 75 VSGREYNIPSVHVGVKSHGWLVEGLSREKABELLLPFGGAPLIRESQTRGYSLSV 134  
 Db 62 VSGREYNIPSVHVGVKSHGWLVEGLSREKABELLLPFGGAPLIRESQTRGYSLSV 121  
 QY 135 RLSPASWDRIHRYHICLDNGWLYISPRLTFFPSLQALVDHYSCLAADDICLLKEPCVLIQ 194  
 Db 122 RLSPASWDRIHRYHICLDNGWLYISPRLTFFPSLQALVDHYSCLAADDICLLKEPCVLIQ 181  
 QY 195 RAGPLPGKDIPLPVTQRTPLNWKELSSLLFSEAAATGEESLLSEGLRSLFYISLND 254  
 Db 182 RAGPLPGKDIPLPVTQRTPLNWKELSSLLFSEAAATGEESLLSEGLRSLFYISLND 241  
 QY 255 AVSLDDA 261  
 Db 242 AVSLDDA 248

RESULT 6  
 AAO15456  
 ID AAO15456 standard; protein; 259 AA.  
 XX  
 XX AAO15456;  
 AC  
 XX  
 XX 03-OCT-2002 (first entry)  
 DT  
 XX  
 XX Mouse modulator of antigen receptor signalling (MARS) protein.  
 DE  
 XX  
 XX Mouse; gene therapy; modulator of antigen receptor signalling; MARS;  
 KW  
 KW tumour suppressor gene; Scr-like adaptor protein; SLAP;

KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 OS Mus sp.  
 PN WO200242452-A2.  
 PD 30-MAY-2002.  
 XX 26-NOV-2001; 2001WO-CA001662.  
 PF 27-NOV-2000; 2000CA-02324663.  
 PR (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 PA Mcglade JC, Loreto MP;  
 PI WPI; 2002-566564/60.  
 XX N-PSDB; AAL44087.  
 XX New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX Claim 7; Fig 1A; 110pp; English.  
 PS The invention comprises the amino acid and coding sequences of modulator  
 XX of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present amino acid sequence represents a mouse MARS  
 CC protein  
 XX Sequence 259 AA;  
 SQ Query Match 76.0%; Score 1028; DB 5; Length 259;  
 Best Local Similarity 79.4%; Pred. No. 4.9e-97;  
 Matches 208; Conservative 16; Mismatches 34; Indels 4; Gaps 3;  
 QY 1 MGSILPSRKSLPSPSLSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 DB 1 MGSILSSSGKT-SFSPSSSGGPDQEPVSMQPERHKVATAVALGSPFAGGPAELSLRLGEPLT 59  
 QY 61 IVSEGDGWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLPLGPNPGGAFLI 120  
 DB 60 IISEDGWTVQVEVSGREYHMSVYVAKVAHGLVEGLSREKAEELLPLGPNPGGAFLI 119  
 QY 121 RESQTRGYSLSVLRSPASWDRIHYRTHCLDNGWLYISPLTTPFSLOALVDHYSELA 180  
 DB 120 RESQTRGYSLSVLRSPASWDRIHYRTHCLDNGWLYISPLTTPFSLOALVDHYSELA 179  
 QY 181 DDICLLKEPCVLQACPLKQIDPLPVTQRTPLNWKELDSSLPSSEA-ATGEESLLSE 239  
 DB 180 DGICCPRECEVQLKGLGFLKQTPPTVPTSSLNWKKLDRSLLEHAPASGEASLLSE 239  
 QY 240 GLRESLSFYISLNDEAVSLDDA 261  
 DB 240 GLRESLSFYISLAED--PLDDA 259  
 RESULT 7  
 AA015458  
 ID AAO15458 standard; protein; 210 AA.  
 XX AA015458;  
 XX 03-OCT-2002 (first entry)  
 DT Mouse modulator of antigen receptor signalling short isoform protein.  
 DE

XX Mouse; gene therapy; modulator of antigen receptor signalling; MARS;  
 KW tumour suppressor gene; Scr-like adaptor protein; SLAP;  
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 XX Mus sp.  
 OS WO200242452-A2.  
 PN 30-MAY-2002.  
 XX 26-NOV-2001; 2001WO-CA001662.  
 PF 27-NOV-2000; 2000CA-02324663.  
 PR (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 PA Mcglade JC, Loreto MP;  
 PI WPI; 2002-566564/60.  
 XX N-PSDB; AAL44090.  
 XX New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX Claim 8; Page 78; 110pp; English.  
 PS The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present amino acid sequence represents a mouse MARS  
 CC protein  
 XX Sequence 210 AA;  
 SQ Query Match 69.1%; Score 934.5; DB 5; Length 210;  
 Best Local Similarity 88.9%; Pred. No. 1.5e-87;  
 Matches 184; Conservative 3; Mismatches 13; Indels 7; Gaps 1;  
 QY 1 MGSILPSRKSLPSPSLSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 DB 1 MGSILPSRKSLPSPSLSSVQGGPVTWEAERSKATAVALGSPFAGGPAELSLRLGEPLT 60  
 QY 61 IVSEGDGWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLPLGPNPGGAFLI 120  
 DB 61 IVSEGDGWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLPLGPNPGGAFLI 120  
 QY 121 RESQTRGYSLSVLRSPASWDRIHYRTHCLDNGWLYISPLTTPFSLOALVDHYSE-- 178  
 DB 121 RESQTRGYSLSVLRSPASWDRIHYRTHCLDNGWLYISPLTTPFSLOALVDHYSEGW 180  
 QY 179 ----LADDICLLKEPCVLQACPLP 200  
 DB 181 PAPQGYTPTCDCAEDTTQLERAGQLP 207  
 RESULT 8  
 ABP64788  
 ID ABP64788 standard; protein; 210 AA.  
 XX ABP64788;  
 XX 25-FEB-2003 (first entry)  
 DT Human protein SEQ ID 448.  
 XX

Human; expressed sequence tag; EST; haematopoietic disorder;  
central nervous system disease; viral infection;  
peripheral nervous system disease; non-healing wound; infectious disease;  
immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
fungal infection; autoimmune disorder; coagulation disorder; neutropenic;  
antiallergic; antiinflammatory; immunosuppressive; neuroprotective;  
cytostatic; haemostatic; virucide; antibacterial; fungicide;  
immunostimulant; cerebroprotective.

Homo sapiens.

WO200259260-A2.

01-AUG-2002.

16-NOV-2001; 2001WO-US042950.

17-NOV-2000; 2000US-00714936.

(HYSE-) HYSEQ INC.

Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA,  
Ren F, Xue AU, Yang Y, Wehrman T, Drmanac RT;

WPI; 2002-590824/63.

N-PSDB; ABO99374.

New isolated polynucleotide, useful in research, diagnostic or  
therapeutic methods, e.g. preventing or treating disorders involving  
aberrant protein expression or biological activity.

Claim 20; SEQ ID NO 448; 394pp; English.

The present invention relates to novel human coding sequences (ABO99268-  
ABO99608) and proteins (ABP6482-ABP65022). The sequences are useful in  
therapeutic, diagnostic and research methods. The polynucleotides may be  
used in the field of molecular biology as hybridisation probes, primers  
for PCR, for chromosome and gene mapping, for the recombinant production  
of protein, or in generation of anti-sense DNA or RNA. The  
polynucleotides are useful in diagnostics as expressed sequence tags  
(ESTs) for identifying expressed genes or for physical mapping of the  
human genome. The proteins may be used as molecular weight markers, or as  
nutritional sources or supplements. The proteins may be used to maintain  
and expand cell population in a totipotent or pluripotent state  
useful for re-engineering damaged or diseased tissues, transplantation,  
manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
polynucleotides and proteins are useful for preventing, treating or  
ameliorating disorders involving aberrant protein expression or  
biological activity, e.g. haematopoietic disorders, central/peripheral  
nervous system diseases, mechanical and traumatic disorders, non-healing  
wounds, immune deficiencies and disorders, infectious diseases caused by  
viral, bacterial or fungal infection, autoimmune disorders, allergic  
reactions and conditions, coagulation disorders, or cancer. The  
polynucleotide sequences of the invention were assembled from ESTs  
isolated mainly by sequencing by hybridisation, and in some cases,  
sequences obtained from one or more public databases. Note: The sequence  
data for this patent did not form part of the printed specification, but  
was obtained in electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences

Sequence 210 AA;

Query Match 68.3%; Score 923.5; DB 5; Length 210;  
Best Local Similarity 87.9%; Pred. No. 2.1e-86;  
Matches 182; Conservative 3; Mismatches 15; Indels 7; Gaps 1;

1 MGSLPSRRKSLPSLSSSVQCGPVTWEAERSKATAVALGSPFAGGPAELSLRLGPLET 60  
1 MGSLPSRRKSLPSLSSSVQCGPVTWEAERSKATAVALGSPFAGGPAELSLRLGPLET 60

61 IVSEGDWMTVLSEVSGREYNTPSVHVGVKSHGWLVEGLSRKABELLLPNCGGAFPLI 120  
61 IVSEGDWMTVLSEVSGREYNTPSVHVGVKSHGWLVEGLSRKABELLLPNCGGAFPLI 120

QY 121 RESQTRRGYSLSVRLSRPASWDRIHYRIHCLDNGWLYISFRLTFPSLQALVDHYSE-- 178  
DB 121 RESQTRRGYSLSVRLSRPASWDRIHYRIHCLDNGWLYISFRLTFPSLQALVDHYSEGW 180  
QY 179 -----LADDICCLLKEPCVLRAGBLP 200  
DB 181 PAPWQGYTPTCDCAEDTTOLERAGQLP 207

RESULT 9  
ABP64107  
ID ABP64107 standard; protein; 159 AA.  
XX  
AC ABP64107;  
XX  
DT 04-NOV-2002 (first entry)  
XX  
DE Human ORF477.  
XX  
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
KW cancer; cardiovascular disease; allergy; autoimmune disease;  
KW wound healing; blood coagulation disorder; inflammatory disorder.  
XX  
OS Homo sapiens.  
XX  
FN US2002082206-A1.  
XX  
PD 27-JUN-2002.  
XX  
PF 30-MAY-2001; 2001US-00867550.  
XX  
PR 30-MAY-2000; 2000US-0208427P.  
PA (LEAC/) LEACH M D.  
PA (MEHR/) MEHRABAN F.  
PA (CONL/) CONLEY P B.  
PA (TOPP/) TOPPER J N.  
PA (LAWD/) LAW D.  
XX  
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
XX  
DR WPI; 2002-626554/67.  
DR N-PSDB; ABO98670.  
XX  
PT New polypeptide designated ORFX are present in human atherogenic cells  
PT and are useful to prevent and treat ORFX-associated disorders including  
PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
PT inflammatory disease.  
XX  
PS Claim 10; SEQ ID NO 954; 78pp; English.  
XX  
CC The present invention relates to novel human ORFX polypeptides and their  
CC coding sequences (ABP63631-ABP64681 and ABO98194-ABO99267). The sequences  
CC were discovered in human atherogenic cells, in particular in platelets  
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
CC many other tissues as well. Atherogenic cells are cells which have the  
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
CC nucleic acids are useful for treating or preventing a pathological  
CC condition associated with an ORFX-associated disorder, e.g. cancer,  
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
CC coagulation disorders or inflammatory disorders. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from the USPTO web site at  
CC seqdata.uspto.gov/sequence.html?DocID=2002082206  
XX  
SQ Sequence 159 AA;

Query Match 61.0%; Score 826; DB 5; Length 159;  
Best Local Similarity 99.4%; Pred. No. 1.5e-76;  
Matches 159; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSLPSRRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRIGELPT 60  
 DB 1 MGSLPSRRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRIGELPT 60  
 QY 61 IVSEDDGWTWTLSEVSGREYNIPSVHGVKSHGLYEGLSREKAEELLPLGPNPGGAFLI 120  
 DB 61 IVSEDDGWTWTLSEVSGREYNIPSVHGVKSHGLYEGLSREKAEELLPLGPNPGGAFLI 120  
 QY 121 RESQTRGSSLSVRLSRPASWDRIHRHRIHCLDNGWLY 159  
 DB 121 RESQTRGSSLSVRLSRPASWDRIHRHRIHCLDNGWLY 159

RESULT 10  
 ABP64588  
 ID ABP64588 standard; protein; 113 AA.  
 XX  
 AC ABP64588;  
 XX  
 DT 04-NOV-2002 (first entry)  
 XX  
 DE Human ORF958.  
 XX  
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnary;  
 KW Antiinflammatory; Gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002082206-A1.  
 XX  
 PD 27-JUN-2002.  
 XX  
 PF 30-MAY-2001; 2001US-00867550.  
 XX  
 PR 30-MAY-2000; 2000US-0208427P.  
 XX  
 PA (LEAC/) LEACH M D.  
 PA (MEHR/) MEHRABAN F.  
 PA (CONL/) CONLEY P B.  
 PA (TOPP/) TOPPER J N.  
 PA (LAWD/) LAW D.  
 XX  
 PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX  
 DR WPI; 2002-626554/67.  
 XX  
 DR N-PSDB; ABQ99151.  
 XX  
 PT New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 XX  
 PS Claim 10; SEQ ID NO 1916; 78pp; English.  
 XX  
 CC The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?DocID=2002082206  
 XX  
 SQ Sequence 113 AA;

Query Match 43.3%; Score 586; DB 5; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-52;  
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 150 IHLCDNGWLYISPRITPSPSQALVDHYSELADICLLKEPCVLQACPLFGKDIPLPVT 209  
 DB 2 IHLCDNGWLYISPRITPSPSQALVDHYSELADICLLKEPCVLQACPLFGKDIPLPVT 61  
 QY 210 VQRTPLNWKELDSSLLFSEATGSESLSEGRLSFLYISLNDEAVSLDDA 261  
 DB 62 VQRTPLNWKELDSSLLFSEATGSESLSEGRLSFLYISLNDEAVSLDDA 113

RESULT 11  
 ABP52188  
 ID ABP52188 standard; protein; 276 AA.  
 XX  
 AC ABP52188;  
 XX  
 DT 15-OCT-2002 (first entry)  
 XX  
 DE Human Src-like adapter protein (SLAP).  
 XX  
 KW Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;  
 KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;  
 KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;  
 KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
 KW chronic inflammatory disorder; autoimmune disorder; transplant rejection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200255707-A2.  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 10-JAN-2002; 2002WO-US000718.  
 XX  
 PR 10-JAN-2001; 2001US-0260953P.  
 XX  
 PA (RIGE-) RIGEL PHARM INC.  
 XX  
 PI Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
 XX  
 DR WPI; 2002-575432/61.  
 XX  
 PT New src-like inhibitory molecule protein, useful for treating  
 PT immunodeficiency disorders and inflammatory disorders, comprises N-  
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.  
 XX  
 PS Example 2; Fig 2B; 91pp; English.

The present invention describes the human Src-like inhibitory molecule (SLIM) protein (I). The present invention describes a SLIM protein comprising an N-terminal myristylation sequence, an N-terminal SH2 domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising an N-terminal myristylation sequence and an N-terminal SH2 domain which is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and anti-HIV activities, and can be used as a modulator of lymphocyte activation, and of ubiquitination of a Cbl target protein, and in gene therapy. (I) is useful for screening a bioactive agent capable of binding to SLIM. (I) is also useful for screening a bioactive agent capable of modulating SLIM binding. (I) or its fragments is useful in the study or in the treatment of conditions which involve this function or dysregulation of SLIM protein activity, i.e. to diagnose, treat or prevent SLIM associated disorders. (I) or the polynucleotide encoding it (II) is useful for modulating leukocyte and/or platelet activation, for modulating antigen receptor-induced signaling and activation in leukocyte and/or platelets and for modulating antigen receptor-induced signaling and activation in lymphocytes and/or mast cells. (I) or (II) is also useful for modulating the basal activity of lymphocytes. (I) or (II) is useful in the treatment of immunodeficiency disorders, such as

CC acquired immunodeficiency syndrome (AIDS), for the prevention and  
 CC treatment of acute inflammatory disorders, chronic inflammatory  
 CC disorders, autoimmune disorder and transplant rejection. The present  
 CC sequence represents the human Src-like adapter protein (SLAP), which is  
 CC given in comparison with the human SLIM protein in an example from the  
 CC present invention  
 XX  
 SQ Sequence 276 AA;  
 Query Match 35.6%; Score 481.5; DB 5; Length 276;  
 Best Local Similarity 40.3%; Pred. No. 1.1e-40;  
 Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;  
 QY 9 KSLPSPSLSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPITIVSEGDW 68  
 Db 6 KSTPAPA-----ERPLNPEGLSDFLAVLSDYSPDISPPIFRGKGLVISEGCGW 58  
 QY 69 WTVLSEVSGREYNIPSVHGVKSHGWLVEG-SREKAEELLLPGNPGGAFLRESQTRRG 128  
 Db 59 WKAISLSTGRESYIPGICVARVYHGWLFEGGLDRKAEELLQPLDTKVGSMFIRESETKG 118  
 QY 129 SYSLSVRLSRPASWDRIRHYRICHLDNGWLVTSPRLTSPSLQALVDHYSELADICCLLK 188  
 Db 119 FYSLSVR-----HRQVKHYRIFRLPNWYIISPLTFCLEDLVNHYSEVADGLCCVLT 172  
 QY 189 EPCVLQRAQGLPGKDIPLVTVQRTPLNWKELDSSLLFSEAATG-----EESLLSEGL 241  
 Db 173 TPCLTQSTAAPAVRASSSPVTLRQKTVDRVRSR---LQEDPEGTENPLGVDESLSFVGL 229  
 QY 242 RESLSFYISLNDE 254  
 Db 230 RESIASYLSLTSE 242  
 RESULT 12  
 ID ABU79103 standard; protein; 276 AA.  
 XX  
 AC ABU79103;  
 XX  
 DT 18-JUN-2003 (first entry)  
 DE Src-like adaptor protein, SLAP.  
 XX  
 KW Superantigen; SAG; staphylococcal enterotoxin; tumour; cancer; apoptosis;  
 KW gene therapy; mammalian cell receptor; cytostatic;  
 KW tumour associated lipid; energy; T cell; antigen presenting cell; APC;  
 KW tumouricidal immunocyte; antitumour.  
 XX  
 OS Unidentified.  
 XX  
 PN US2002177551-A1.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 30-MAY-2001; 2001US-00870759.  
 XX  
 PR 31-MAY-2000; 2000US-0208128P.  
 XX  
 PA (TERM/) TERMAN D S.  
 XX  
 PI Terman DS;  
 XX  
 PI WPI; 2003-361759/34.  
 XX  
 DR A mammalian cell receptor, useful in the treatment of cancer by binding  
 PT to tumor associated lipids where the binding induces energy or apoptosis  
 PT in T cells and antigen presenting cells.  
 XX  
 PS Disclosure; Page; 167pp; English.  
 XX  
 CC The invention relates to a mammalian cell receptor, useful in the  
 CC treatment of cancer, which binds to tumour associated lipids and induces

CC energy or apoptosis in the T cells and antigen presenting cells (APCs).  
 CC Also included are a mammalian cell useful in the treatment of cancer  
 CC where the receptor which binds tumour associated lipids and induces  
 CC cellular inactivation or death is deleted or functionally deactivated,  
 CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal  
 CC (by allowing tumour associated lipids to contact immunocytes in which  
 CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,  
 CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,  
 CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or  
 CC deleted), a construct useful in the treatment of cancer comprising a  
 CC superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell  
 CC useful in the treatment of cancer (where an adaptor protein which  
 CC inhibits T cell activation by tumour associated antigens is deleted or  
 CC functionally deactivated), a composition useful in the treatment of  
 CC cancer (comprising a lipid raft conjugated to a superantigen), producing  
 CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by  
 CC allowing tumour associated lipids to contact immunocytes, in which  
 CC receptors for the lipids are inactivated or deleted to produce a  
 CC tumouricidal immunocyte population, and administering the tumouricidal  
 CC activated immunocytes to the host), producing (M3) a tumouricidal APC  
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to  
 CC contact APCs, in which receptors for the tumour associated lipids are  
 CC inactivated or deleted to produce a tumouricidally activated population,  
 CC and administering APCs to the host), producing a tumouricidal T cell  
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to  
 CC contact T cells, in which adaptor proteins, which inhibit T cell  
 CC activation by tumour associated antigens, are deleted or functionally  
 CC deactivated to produce a tumouricidal population of T cells, and  
 CC administering the tumouricidally activated T cells to the host, or  
 CC allowing a superantigen-lipid raft to contact T cells ex vivo, and  
 CC administering the tumouricidally activated T cells to the host), treating  
 CC (M5) cancer in a mammal (by administering a lipid binding molecule which  
 CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)  
 CC a tumouricidal T cell population in vivo in a mammal (by allowing a  
 CC tumour associated antigen to contact immunocytes in which adaptor  
 CC proteins which inhibit T cell activation by tumour associated antigens  
 CC are deleted or functionally deactivated) and producing (M7) a  
 CC tumouricidal T cell population ex vivo in a mammal comprising allowing a  
 CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The  
 CC receptors, methods and compositions are useful for treating cancers and  
 CC tumours. Bacterial superantigens are co-administered or administered as  
 CC fusion constructs with anti-tumour proteins or motifs. The present  
 CC sequence represents a tumour antigen or a motif identifying a tumour  
 CC antigen, which can be functionally deactivated in the method of the  
 CC invention. Note: the sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format from the  
 CC US patent office website at  
 CC "seqdata.uspto.gov/sequence.html?DocID=20020177551"  
 XX  
 SQ Sequence 276 AA;

Query Match 35.6%; Score 481.5; DB 6; Length 276;  
 Best Local Similarity 40.3%; Pred. No. 1.1e-40;  
 Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;  
 QY 9 KSLPSPSLSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPITIVSEGDW 68  
 Db 6 KSTPAPA-----ERPLNPEGLSDFLAVLSDYSPDISPPIFRGKGLVISEGCGW 58  
 QY 69 WTVLSEVSGREYNIPSVHGVKSHGWLVEG-SREKAEELLLPGNPGGAFLRESQTRRG 128  
 Db 59 WKAISLSTGRESYIPGICVARVYHGWLFEGGLDRKAEELLQPLDTKVGSMFIRESETKG 118  
 QY 129 SYSLSVRLSRPASWDRIRHYRICHLDNGWLVTSPRLTSPSLQALVDHYSELADICCLLK 188  
 Db 119 FYSLSVR-----HRQVKHYRIFRLPNWYIISPLTFCLEDLVNHYSEVADGLCCVLT 172  
 QY 189 EPCVLQRAQGLPGKDIPLVTVQRTPLNWKELDSSLLFSEAATG-----EESLLSEGL 241  
 Db 173 TPCLTQSTAAPAVRASSSPVTLRQKTVDRVRSR---LQEDPEGTENPLGVDESLSFVGL 229  
 QY 242 RESLSFYISLNDE 254  
 ||||:|||||  
 ||||:|||||

Db 230 RESIASYLSLTSE 242

RESULT 13

ABR59698

ID ABR59698 standard; protein; 276 AA.

AC ABR59698;

XX

DT 25-JUL-2003 (first entry)

XX

DE Human Src-like adaptor.

XX

XX Human; T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; asthma;

KW immunosuppressive; antiaesthatic; anti-allergic; anti-inflammatory;

KW lymphocyte activation; lymphocyte migration; cytokine production;

KW cell surface marker expression; antibody production; apoptosis; allergy;

KW antibody proliferation; antibody differentiation; hypersensitivity;

KW graft versus host disease; inflammation; Src-like-adaptor.

XX

OS Homo sapiens.

XX

XX WO2003029277-A2.

PN

XX

PD 10-APR-2003.

XX

PF 02-OCT-2002; 2002WO-US031618.

XX

PR 03-OCT-2001; 2001US-0327212P.

XX

PA (RIGB-) RIGEL PHARM INC.

XX

PI Chu P, Li C, Liao XC, Masuda E, Pardo J, Zhao H;

PI

DR WPI; 2003-363276/34.

XX

DR N-PSDB; ACC81091.

XX

XX Identifying a compound that modulates T lymphocyte activation, useful for

PT monitoring changes in cell surface marker expression, comprises

PT contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with

PT a compound.

PT

XX Disclosure; Page 72; 126pp; English.

PS

XX

XX The invention relates to a novel method for identifying a compound that

CC modulates T lymphocyte activation. The method comprises contacting a T

CC cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,

CC where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic

CC acid that hybridises to a nucleic acid encoding a polypeptide having a

CC sequence selected from two 606-amino acid sequence and a 415-amino acid

CC sequence given in the specification. The method of the invention has

CC immunosuppressive, antiaesthatic, anti-allergic, and anti-inflammatory

CC activity. The method is useful for identifying compounds that modulate

CC lymphocyte activation and migration, and for monitoring changes in cell

CC surface marker expression, cytokine production, antibody production,

CC proliferation and differentiation, and apoptosis, using either cell lines

CC or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as

CC drug targets for compounds that suppress or activate lymphocyte

CC activation and migration, e.g. for the treatment of diseases in which

CC modulation of the immune response is desired such as delayed type

CC hypersensitivity reactions, asthma, allergies, graft versus host disease,

CC and acute and chronic inflammation. Modulators of lymphocyte activation

CC are useful for treating disorders related T and B cell activation and

CC migration. The present sequence is used in the exemplification of the

CC invention

XX

XX Sequence 276 AA;

SQ

Query Match 35.6%; Score 481.5; DB 6; Length 276;

Best Local Similarity 40.3%; Pred. No. 1.1e-40;

Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;

QY 9 KSLPSPSLSSVQGGPVTWEABRSKATVALGSPFAGGPAELSLNLGEPPLTVSEDDGW 68

Db 6 KSTPAPA-----ERLPNPEGLSDFLVLSYSPDISPPIFRGKLRVSDGGW 58

QY 69 WTVLSEVSGREYNIPSVHGVKSHGWLVEGLSREKABELLLPGNPGGAFILRESQTRRG 128

Db 59 WKAISSLTGRESYIPGICVARVYHGLFEGLRDKABELLQLPDTKVGSFMRSETKKG 118

QY 129 SYSLSVRLSRPASWDRIHRIHCLDNGWLVIYSPRLTFPSLQALVDHYSELADICLLK 188

Db 119 FYLSLSVR-----HRQVKHYRIFELPNNWYIYSPRLTFQCLDINWYSEVADGLCCVLT 172

QY 189 EPCVLQRAGPLPGKDIPLPVTQVOTPLNWKELDSLLFSBAATG-----EESLLSEGL 241

Db 173 TPCLTQSTAAPAVRASSPVTLRKTVDMRVSRLQEDPEGTENPLGVDELSFVGL 229

QY 242 RESLSFYISLUNDE 254

Db 230 RESIASYLSLTSE 242

RESULT 14

AAU31072

ID AAU31072 standard; protein; 315 AA.

XX

AC AAU31072;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secreted protein #1563.

XX

KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX

OS Homo sapiens.

XX

PN WO200179449-A2.

XX

PD 25-OCT-2001.

XX

PF 16-APR-2001; 2001WO-US008656.

XX

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-611725/70.

XX

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

PT vaccination, testing and therapy.

XX

PS Claim 20; Page 399; 765pp; English.

XX

CC The invention relates to novel human secreted polypeptides. The

CC polypeptides and antibodies to the polypeptides are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for

CC identifying agents (agonists and antagonists) that bind to them. Cells

CC expressing the proteins are useful for identifying a therapeutic agent

CC for use in treatment of a pathology related to aberrant expression or

CC physiological interactions of the polypeptide. Vectors comprising the

CC nucleic acids encoding the polypeptides and cells genetically engineered

CC to express them are also useful for producing the proteins. The proteins

CC are useful in genetic vaccination, testing and therapy, and can be used

CC as nutritional supplements. They may be used to increase stem cell

CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon

CC and/or nerve tissue growth or regeneration; immune suppression and/or

CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.

CC AAU29510-AAU33304 represent the amino acid sequences of novel human

CC secreted proteins of the invention



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:19:42 ; Search time 45 Seconds

(without alignments)  
557.911 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MSLPSPRRKSLSPSPSSSV.....RESLSFYSLNDEAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	493.5	36.5	281	2 A57152	src-like adaptor p
2	370.5	27.4	512	1 TVHULY	protein-tyrosine k
3	367.5	27.2	512	1 I56160	protein-tyrosine k
4	367.5	27.2	512	1 A39719	protein-tyrosine k
5	360.5	26.6	505	1 TVHUHC	protein-tyrosine k
6	353.5	26.1	503	1 JQ1321	protein-tyrosine k
7	352.5	26.1	505	2 I37206	protein-tyrosine k
8	352	26.0	503	1 TVMSHC	protein-tyrosine k
9	344.5	25.5	499	1 A40092	protein-tyrosine k
10	340	25.1	509	1 OKHULK	protein-tyrosine k
11	338.5	25.0	509	1 I48845	protein-tyrosine k
12	333	24.6	507	1 A39939	protein-tyrosine k
13	325	24.0	539	2 B49114	protein-tyrosine k
14	320	23.7	528	1 TVFVG9	protein-tyrosine k
15	319.5	23.6	537	1 A45501	protein-tyrosine k
16	317	23.4	541	1 TVCHYS	protein-tyrosine k
17	315.5	23.3	537	1 TVHUSY	protein-tyrosine k
18	315.5	23.3	543	1 TVHUSY	protein-tyrosine k
19	312.5	23.1	529	1 TVHUFR	protein-tyrosine k
20	311	23.0	544	2 I51593	protein-tyrosine k
21	310.5	22.9	537	1 A43806	protein-tyrosine k
22	309	22.8	534	1 A44991	protein-tyrosine k
23	307	22.7	517	2 S24547	protein-tyrosine k
24	306.5	22.7	541	2 S31645	protein-tyrosine k
25	305.5	22.6	537	2 I51592	protein-tyrosine k
26	301	22.2	517	2 A43807	protein-tyrosine k
27	301	22.2	536	2 S33569	protein-tyrosine k
28	297.5	22.0	534	1 S33568	protein-tyrosine k
29	295.5	21.8	542	2 A49114	protein-tyrosine k

30	290	21.4	506	1 S24553	protein-tyrosine k
31	286.5	21.2	557	1 TVFVS2	protein-tyrosine k
32	286.5	21.2	587	1 TVFVPR	protein-tyrosine k
33	284.5	21.0	533	1 TVCHS	protein-tyrosine k
34	284.5	21.0	568	1 TVFVS1	protein-tyrosine k
35	279.5	20.7	526	1 TVFV60	protein-tyrosine k
36	276	20.4	509	1 TVHAST	protein-tyrosine k
37	275	20.3	546	2 S52314	protein-tyrosine k
38	274.5	20.3	532	1 B34104	protein-tyrosine k
39	273.5	20.2	526	1 TVFVR	protein-tyrosine k
40	273.5	20.2	526	2 S15582	protein-tyrosine k
41	273.5	20.2	542	1 TVHUSC	protein-tyrosine k
42	273	20.2	545	2 S52313	protein-tyrosine k
43	272	20.1	541	1 A43610	protein-tyrosine k
44	271.5	20.1	532	1 A34104	protein-tyrosine k
45	269.5	19.9	526	2 S26420	protein-tyrosine k

#### ALIGNMENTS

##### RESULT 1

A57152

src-like adaptor protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 08-Dec-1995 #sequence\_revision 08-Dec-1995 #text\_change 12-Feb-1999

C:Accession: A57152

R: Pandey, A.; Duan, H.; Dixit, V.M.

J. Biol. Chem. 270, 19201-19204, 1995

A:Title: Characterization of a novel Src-like adapter protein that associates with the R

A:Reference number: A57152; MUID:95370243; PMID:7543898

A:Accession: A57152

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-261 <PAN>

A:Cross-references: GB:U29056

C:Superfamily: SH3 homology; SH2 homology

F:29-77/Domain: SH3 homology <SH3>

F:84-175/Domain: SH2 homology <SH2>

Query Match 36.5%; Score 493.5; DB 2; Length 281;

Best Local Similarity 43.6%; Pred. No. 2.1e-33;

Matches 115; Conservative 36; Mismatches 82; Indels 29; Gaps 7;

QY	9	KSLPSPS---	LSSSVQCGPVTMEASR	KATAVALGFFPAGGPAELSRLGELPTIV	SESD 65
DB	6	KSTSPSPSERPLSSS-	-----	EGLESDFLAV-LTDYPSSDISPPIFRGKELR	VSDE 55
QY	66	GDWMTVLSEVSGREYNT	PSVHVGVKSHGLYEGLSREKAE	LELLLPNPGGAPLIRE	SQT 125
DB	56	GGWKAISLSTGREGSY	IPGICVARVYHGLFEG	LGRDKAEELQLPDTKIGSPMIRE	SET 115
QY	126	RRGSYSLSVRLSRDPAS	WDRIRHRIHCLDNGWLYIS	PRLPSPSLQALVDHYSELAD	ICC 185
DB	116	KKGYLSLSVR-----	HRQVKHYRIFELPNWY	YISRLTFQCLEDLVTHYSE	VADGLCC 169
QY	186	LLKEPCVQLR-----	AGPLGCKDIPLVTVQRT	PLNWKELDSSLLFSEATG	-----BESL 236
DB	170	VLTPFCLAQIPAPTSH	SPSPCTSPGSPVTLRQK	TDFDKRVSRLOQEGSEGA	ENPLRVDES 229
QY	237	LSEGLRESLSFYISL	-NDEAVSLD 259		
DB	230	PSYGLRESIASYLS	LTGDDSSPD 253		

##### RESULT 2

TVHULY

protein-tyrosine kinase (EC 2.7.1.112) lyn, splice form A - human

N:Contains: protein-tyrosine kinase lyn, splice form B

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 18-Feb-2000

C:Accession: A26719; D38268; P09499; I53715

R:Yamanashi, Y.; Fukushima, S.I.; Semba, K.; Sukegawa, J.; Miyajima, N.; Matsubara, K.;

Mol. Cell. Biol. 7, 237-243, 1987  
 A;Title: The yes-related cellular gene lyn encodes a possible tyrosine kinase similar to  
 A;Reference number: A26719; MUID:87172710; PMID:3561390  
 A;Accession: A26719  
 A;Molecule type: mRNA  
 A;Residues: 1-512 <YAM>  
 A;Cross-references: GB:M16038; NID:q187268; PIDN:AA859540.1; PID:q307144  
 R;Paranen, J.; Maekela, T.P.; Alicata, R.; Lehtvaeslaho, H.; Alitalo, K.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990  
 A;Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.  
 A;Reference number: A39268; MUID:91062389; PMID:2247464  
 A;Accession: D38268  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 369-424 <PAR>  
 R;Bieleke, W.; Ziemekki, A.; Kappos, L.; Miescher, G.C.  
 Biochem. Biophys. Res. Commun. 186, 1403-1409, 1992  
 A;Title: Expression of the B cell-associated tyrosine kinase gene lyn in primary neurobl  
 A;Reference number: PH0949; MUID:92378604; PMID:1510669  
 A;Accession: PH0949  
 A;Molecule type: mRNA  
 A;Residues: 369-424 <BIE>  
 A;Experimental source: neuroblastoma SK-IN cell  
 R;Rider, L.G.; Raben, N.; Miller, L.; Jelsema, C.  
 Gene 138, 219-222, 1994  
 A;Title: The cDNAs encoding two forms of the LYN protein tyrosine kinase are expressed i  
 A;Reference number: I53715; MUID:94171041; PMID:8125304  
 A;Accession: I53715  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-24,46-512 <RID>  
 A;Cross-references: GB:M79321; NID:q187270; PIDN:AA850019.1; PID:g187271  
 A;Experimental source: splice form B  
 C;Genetics:  
 A;Gene: GDB:LYN  
 A;Cross-references: GDB:120159; OMIM:165120  
 A;Map position: 8q13-8qter  
 C;Function:  
 A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
 C;Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprote  
 yrosine-specific protein kinase  
 F;2-512/Product: protein-tyrosine kinase  
 F;2-24,46-512/Product: protein-tyrosine kinase lyn, splice form A #status predicted <MATA>  
 F;70-118/Domain: SH3 homology <SH3>  
 F;129-226/Domain: SH2 homology <SH2>  
 F;245-504/Domain: protein kinase homology <KIN>  
 F;253-261/Region: protein kinase ATP-binding motif  
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F;3/Binding site: palmitate (Cys) (covalent) #status predicted  
 F;275/Active site: Lys #status predicted  
 F;397,508/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

RESULT 3  
 I56160  
 N;Contains: protein-tyrosine kinase lyn, splice form A - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 18-Feb-2000 #sequence revision 18-Feb-2000 #text\_change 18-Feb-2000  
 C;Accession: I56160; I67811; I67812  
 R;Minoguchi, K.; Nishikata, H.; Siraganian, R.P.  
 J. Immunol. 150, 222, 1993  
 A;Title: Bacterially expressed rat p56lyn binds several proteins in rat basophilic leuke  
 A;Reference number: I56160  
 A;Accession: I56160  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-512 <MIN>  
 A;Cross-references: GB:I14951; NID:q294582; PIDN:AAA41549.1; PID:q294583  
 R;Rider, L.G.; Raben, N.; Miller, L.; Jelsema, C.  
 Gene 138, 219-222, 1994  
 A;Title: The cDNAs encoding two forms of the LYN protein tyrosine kinase are expressed i  
 A;Reference number: I53715; MUID:94171041; PMID:8125304  
 A;Accession: I67811  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-230, 'L', 232-307, 'A', 309-418, 'Y', 420-512 <RID1>  
 A;Cross-references: GB:I14782; NID:q294578; PIDN:AAA20944.1; PID:q294579  
 A;Note: in Genbank entry RATLYNATYR, release 116.0, PIDN:AAA20944.1, the source is desig  
 A;Accession: I67812  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-24,46-230, 'L', 232-307, 'A', 309-418, 'Y', 420-512 <RID2>  
 A;Cross-references: GB:I14823; NID:q294580; PIDN:AAA20945.1; PID:q294581  
 A;Note: in Genbank entry RATLYNATYR, release 116.0, PIDN:AAA20945.1, the source is desig  
 C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
 C;Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprote  
 F;2-512/Product: protein-tyrosine kinase lyn, splice form A #status predicted <MATA>  
 F;2-24,46-512/Product: protein-tyrosine kinase lyn, splice form B #status predicted <MAT  
 F;70-118/Domain: SH3 homology <SH3>  
 F;129-226/Domain: SH2 homology <SH2>  
 F;245-504/Domain: protein kinase homology <KIN>  
 F;253-261/Region: protein kinase ATP-binding motif  
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F;275/Active site: Lys #status predicted  
 F;397,508/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 27.2%; Score 367.5; DB 1; Length 512;  
 Best Local Similarity 40.3%; Pred. NO. 1.2e-22;  
 Matches 81; Conservative 33; Mismatches 78; Indels 9; Gaps 3;  
 Qy 6 SRKSLPSPSLSSVGGPVTWEAERSKATAVALGSPAGGPAELSLRLGEPITVSED 65  
 Db 38 SNKQRPVPE-SQLLPQGRFQKDPPEEQDIVVLPYDGIHPDDLSFKKGKMKVLEH 96  
 Qy 66 GDMWTVLSEVSGREYNIPSVHVGK-----SHGWLVEGLSREKAEELLLPGNPGGAFLIR 121  
 Db 97 GEWWAKSLTKKEGFIIPSNYAKVNTLEETEEFFKDIKDAERQLLAPGNSAGAFILIR 156  
 Qy 122 ESOTRSGVSLVRLSRPASWRIHYRHCLDNGWLVIYSPRLTPPSLQALVDHYSELAD 181  
 Db 157 ESETLKGSPSLVRDPPVHGDVVKHYKIRSLDNGYIYSPRITFPFCISDMIKHYKQSD 216  
 Qy 182 DICLLKPCVQLQRAQPLPK 202  
 Db 217 GLCRLKAKI-----SPKPK 233

RESULT 4  
 A39719  
 N;Contains: protein-tyrosine kinase (EC 2.7.1.112) lyn, long splice form - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 18-Feb-2000 #sequence revision 18-Feb-2000 #text\_change 03-Mar-2000  
 C;Accession: A39719; B39719; A39750; B39750

R;Stanley, E.; Ralph, S.; McEwen, S.; Boulet, I.; Holtzman, D.A.; Lock, P.; Dunn, A.R.  
Mol. Cell. Biol. 11, 3399-3406, 1991  
A;Title: Alternatively spliced murine lyn mRNAs encode distinct proteins.  
A;Reference number: A39719; MUID:91260688; PMID:1710766  
A;Accession: A39719  
A;Molecule type: mRNA  
A;Residues: 1-512 <STAL>  
A;Cross-references: GB:M64608; NID:g198938; PIDN:AAA39470.1; PID:g198939  
A;Accession: B39719  
A;Molecule type: mRNA  
A;Residues: 1-24,46-512 <STAL2>  
A;Cross-references: GB:M64608  
R;Yi, T.; Bolen, J.B.; Ihle, J.N.  
Mol. Cell. Biol. 11, 2391-2398, 1991  
A;Title: Hematopoietic cells express two forms of lyn kinase differing by 21 amino acids  
A;Reference number: A39750; MUID:91203857; PMID:2017160  
A;Accession: A39750  
A;Molecule type: mRNA  
A;Residues: 1-76,'F',78-160,'I',162-278,'L',280-390,'I',392-424,'D',426-512 <YII>  
A;Cross-references: GB:M57696; NID:g198940; PIDN:AAA39471.1; PID:g198941  
A;Accession: B39750  
A;Molecule type: mRNA  
A;Residues: 1-24,46-75,'F',78-160,'I',162-278,'L',280-390,'I',392-424,'D',426-512 <YI2>  
A;Cross-references: GB:M57697; NID:g198942; PIDN:AAA39472.1; PID:g198943  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprote  
F;1-512/Product: protein-tyrosine kinase lyn, long splice form #status predicted <MATL>  
F;1-24,46-512/Product: protein-tyrosine kinase lyn, short splice form #status predicted  
F;70-118/Domain: SH3 homology <SH3>  
F;129-226/Domain: SH2 homology <SH2>  
F;245-504/Domain: protein kinase homology <KIN>  
F;253-261/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;275/Active site: Lys #status predicted  
F;397,508/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 27.2%; Score 367.5; DB 1; Length 512;  
Best Local Similarity 40.3%; Pred. No. 1.2e-22;  
Matches 81; Conservative 33; Mismatches 78; Indels 9; Gaps 3;  
QY 6 SRRKSLPSPSSSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPLTIVSED 65  
Db 38 SNKQRPVPPEF-HLLPGQRFQTKPEEGDIVVALYFDGHPDDLSFKKGGKMKVLEEH 96  
QY 66 GDWTVLSEVSGREYNIPSVHGVK-----SHGWLYEGLSRKABELLLPNCPPGAPLIR 121  
Db 97 GEMWKALESKREGFIPSNVAKYNTLETBEWFKDITRKDAERQLLAPGNSAGAPLIR 156  
QY 122 ESQTRRGYSLSVRLSRPASWDRIHRYHICLDNGWLYISPRITFPPSLQALVDHYSLEAD 181  
Db 157 ESETLKGSFSLVSRDYDPMHGDVKKYKIRSLDNGGYIISPRITFPCISDMKHQKQSD 216  
QY 182 DICLLKEPCVLRAGPLPGK 202  
Db 217 GLCRRLEXACI-----SPKPK 233  
RESULT 5  
TVRHUC  
protein-tyrosine kinase (EC 2.7.1.112) hck - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1989 #sequence revision 10-Nov-1995 #text\_change 11-Jun-1999  
C;Accession: A27811; A27812; JC1149; C38268; S31103  
R;Quintrell, N.; Lebo, R.; Varmus, H.; Bishop, J.M.; Pettinati, M.J.; Le Beau, M.M.; Dia  
Mol. Cell. Biol. 7, 2267-2275, 1987  
A;Title: Identification of a human gene (HCK) that encodes a protein-tyrosine kinase and  
A;Reference number: A27811; MUID:87257942; PMID:3496523  
A;Accession: A27811  
A;Molecule type: mRNA  
A;Residues: 1-505 <QUT>  
A;Cross-references: GB:M16591  
A;Note: the codon given for 3-Cys (TCG) is inconsistent with the authors' translation  
R;Ziegler, S.F.; Marth, J.D.; Lewis, D.B.; Perlmutter, R.M.

Mol. Cell. Biol. 7, 2276-2285, 1987  
A;Title: Novel protein-tyrosine kinase gene (hck) preferentially expressed in cells of he  
A;Reference number: A27812; MUID:87257943; PMID:3453117  
A;Accession: A27812  
A;Molecule type: mRNA  
A;Residues: 1-505 <ZIE>  
A;Cross-references: GB:M16592; NID:g183913; PIDN:AAA52644.1; PID:g306833  
R;Radetzky, D.; Strebhardt, K.; Ruebsamen-Waigmann, H.  
Gene 113, 275-280, 1992  
A;Title: The genomic locus of the human hemopoietic-specific cell protein tyrosine kinase  
A;Reference number: JC1149; MUID:92241680; PMID:1572549  
A;Accession: JC1149  
A;Molecule type: DNA  
A;Residues: 157-505 <HRA>  
A;Cross-references: EMBL:X59741  
R;Parranen, J.; Maekela, T.P.; Alitalo, R.; Lehtvaeslahti, H.; Alitalo, K.  
Proc. Natl. Acad. Sci. U.S.A. 87, 8913-8917, 1990  
A;Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.  
A;Reference number: A38268; MUID:91062389; PMID:2247464  
A;Accession: C38268  
A;Status: nucleic acid sequence not shown; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 362-417 <PAR>  
C;Genetics:  
A;Gene: GDB:HCK  
A;Cross-references: GDB:119303; OMIM:142370  
A;Map position: 20q11-20q12  
A;Introns: 207/1; 258/1; 318/1; 343/3; 395/1; 439/1  
C;Function:  
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phos  
F;2-505/Product: protein-tyrosine kinase hck #status predicted <MAT>  
F;64-112/Domain: SH3 homology <SH3>  
F;123-220/Domain: SH2 homology <SH2>  
F;239-497/Domain: protein kinase homology <KIN>  
F;247-255/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;3/Binding site: palmitate (Cys) (covalent) #status predicted  
F;269/Active site: Lys #status predicted  
F;390/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 26.6%; Score 360.5; DB 1; Length 505;  
Best Local Similarity 41.6%; Pred. No. 4.6e-22;  
Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
QY 12 PPSPLSSSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGEPLTIVSEDGDMTV 71  
Db 40 PGPNGHNS--NTPGLREAGSDDIIVALYDYEATHHEDLSFKQGDQMVVLESGEWMKA 96  
QY 72 LSEVSGREYNIPSVHGVK-----SHGWLYEGLSRKABELLLPNCPPGAPLIRSQTR 127  
Db 97 RSLATRKKEGYPISNYVARVDSLETBEWFKGISRKDAERQLLAPGNGMLGSPMIRDSSETTK 156  
QY 128 GSYLSVRLSRPASWDRIHRYHICLDNGWLYISPRITFPPSLQALVDHYSLEADICLL 187  
Db 157 GSYLSVSRDYDPRQGTGVKHYKIRTLNDGGFVIFSRFTSTLQELVDHYHKGNDGLCQK 216  
QY 188 KEPCV 192  
Db 217 SVPCM 221  
RESULT 6  
JQJ321  
protein-tyrosine kinase (EC 2.7.1.112) hck - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 04-Feb-2000  
C;Accession: JQJ321; S18974  
R;Okano, Y.; Sugimoto, Y.; Fukuoka, M.; Matsui, A.; Nagata, K.; Nozawa, Y.  
Biochem. Biophys. Res. Commun. 181, 1137-1144, 1991  
A;Title: Identification of rat cDNA encoding hck tyrosine kinase from megakaryocytes.  
A;Reference number: JQJ321; MUID:92109719; PMID:1764064

```

F,2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F,269/Active site: Lys #status predicted

Query Match          26.1%; Score 352.5; DB 2; Length 505;
Best Local Similarity 43.6%; Pred. No. 2.1e-21;
Matches 75; Conservative 24; Mismatches 68; Indels 5; Gaps 2;

QY 25 PVTMEARSKATAVALGSPAGGPAELSLRLGEPLTIVSEDDWMTVLSEVSGREYNIPS 84
DB 51 PDEHLDEKHFVVALYDYTAAMDRLQMLAKGKLVQLKGTGDWMLARSLVTGREGVPS 110

QY 85 VHVGVKVS-----HGWYEGLSREKABEILLPLGNPGGAPLIRESTRGYSYLSVLSRPA 140
DB 111 NVPAVSEVLEMERWFRFSQGRKEARQLLAPINKAGSPLIRESETNKGAFSLSVK-DVTT 169

QY 141 SWDRIRHYRIHCLDNGWLYISPRITFSPSLQALVDHYHSELADDICLLKEPCV 192
DB 170 QCELIKHYKIRCLDEGGVYISPRITFSPSLQALVQHYSKGDLGCLORULTPCV 221

RESULT 8
TVMSHC
N;protein-tyrosine kinase (EC 2.7.1.112) hck - mouse
N;Alternate names: kinase-related transforming protein (bmk)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 28-Jan-2000
C;Accession: A27282, A39973
R;Klemsz, W.J.; McRercher, S.R.; Maki, R.A.
Nucleic Acids Res. 15, 9600, 1987
A;Title: Nucleotide sequence of the mouse hck gene.
A;Reference number: A27282; MUID:88067781; PMID:3684607
A;Accession: A27282
A;Molecule type: mRNA
A;Residues: 1-503 <LE>
A;Cross-references: GB:Y00487; NID:951209; PIDN:CAA68544.1; PID:951210
R;Holtzman, D.A.; Cook, W.D.; Dunn, A.R.
Proc. Natl. Acad. Sci. U.S.A. 84, 8325-8329, 1987
A;Title: Isolation and sequence of a cDNA corresponding to a src-related gene expressed
A;Reference number: A39973; MUID:88068587; PMID:3317404
A;Accession: A39973
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-503 <HOL>
A;Cross-references: GB:J03023; NID:g192212; PIDN:AAA37305.1; PID:g309118
C;Genetics: hck
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology; SH3 hck
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phospho
F;62-110/Domain: SH3 homology <SH3>
F;121-218/Domain: SH2 homology <SH2>
F;237-495/Domain: protein kinase homology <KIN>
F;245-253/Region: protein kinase ATP-binding motif
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;3/Binding site: palmitate (Cys) (covalent) #status predicted
F;267/Active site: Lys #status predicted
F;388,499/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match          26.0%; Score 352; DB 1; Length 503;
Best Local Similarity 41.0%; Pred. No. 2.3e-21;
Matches 82; Conservative 33; Mismatches 73; Indels 12; Gaps 4;

QY 11 LPSPLSSVQCGQPVTE-----AERSKAT-AVALGSPAGGPAELSLRLGFLTIVSEDG 66
DB 30 VDPFTSSSKLGPNNSNPPGFVEGSEDTIVVALYDYEAHREDLSFGQDQMVVLEAG 89

QY 67 DWMTVLSEVSGREYNIPSVHGVK-----SHGWYEGLSREKABEILLPLGNPGGAFIRE 122
DB 90 EWWKARSLATKKEGYIPSNYVARVNSLTEEWFVKIGSRKDAERHLLAPGNMLGSFMIRD 149

QY 123 SQTREGSYLSVLSRPAASWDRIRHYRTHCLDNGWLYISPRITFSPSLQALVDHYHSELADD 182
DB 150 SETTKGYSLSVDRDPDQGHDTIVKRYKIRTLDSGGFYISPRSTSSQLDELVLHYKKGKD 209

```

QY 183 ICCLKEPCVQLQKAGPLGK 202  
Db 210 LCQKLSVPCV----SPKPK 225

RESULT 9  
A40092  
protein-tyrosine kinase (EC 2.7.1.112) blk [validated] - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 16-Jun-2000 #sequence\_revision 16-Jun-2000 #text\_change 16-Jun-2000  
C;Accession: A40092  
R;Dymekci, S.M.; Niederehuber, J.E.; Desiderio, S.V.  
Science 247, 332-336, 1990  
A;Title: Specific expression of a tyrosine kinase gene, blk, in B lymphoid cells.  
A;Reference number: A40092; MUID:90117147; PMID:2404338  
A;Accession: A40092  
A;Molecule type: mRNA  
A;Residues: 1-499 <DYM>  
A;Cross-references: GB:M30903; NID:G202076; PIDN:AAA40453.1; PID:G202077  
C;Genetics:  
A;Gene: MGI:Blk  
A;Map position: 14:26.0  
A;Cross-references: MGI:88169  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology  
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho  
F;59-107/Domain: SH2 homology <SH2>  
F;118-214/Domain: SH2 homology <SH2>  
F;233-491/Domain: protein kinase homology <KIN>  
F;241-249/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;263/Active site: Lys #status predicted

Query Match 25.5%; Score 344.5; DB 1; Length 499;  
Best Local Similarity 36.0%; Pred. No. 9.5e-21;  
Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;

QY 1 MGSLPERRKSLPSLSVSVQGGPV-----TWAEERSK 34  
Db 1 MGLLSKRRQ-----VSEKGGKMSPKINTQKAPPLPLVFNHLAPPGNQDPDEE 54

QY 35 ATAVALGSPFAGGPAELSLRLGEPLTIVSEDDGWTVLSEVSGREYNIPSVHVGKVS-- 91  
Db 55 RFVVALFDYAANDRLQVLKKEKLVLRSTGDMWLARSLVTRGGYVPSNFVAPVEILE 114

QY 92 -HGWLYEGLSRKAEELLLLPNPGGAFIRSQTRRSYSLSVLSRSPASDRIRHRYI 150  
Db 115 VEKMFPTTISRKDAERQLLAPNKKAGSFLIRSESNNKGAFSLSVK-DITTOGEVVKHYKI 173

QY 151 HCLDNGMLVSLRPLTSPSQALVDHYSELADDCILKEPCV 192  
Db 174 RSLDNGGYVISRPTFTLQALVQHYSKGDLCKQLTLCV 215

RESULT 10  
ORHULK  
N;Alternate names: kinase-related transforming protein (lck)  
C;Species: Homo sapiens (man)  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 28-Jan-2000  
C;Accession: JQ0152; S07822; S07200; S01879; S07143; A32797; I57636  
R;Rouer, E.; Van Huynh, T.; de Souza, S.L.; Lang, M.C.; Fischer, S.; Benarous, R.  
Gene 84, 105-113, 1989  
A;Title: Structure of the human lck gene: differences in genomic organisation within src  
A;Reference number: JQ0152; MUID:90108697; PMID:2558056  
A;Accession: JQ0152  
A;Molecule type: DNA  
A;Residues: 1-509 <ROU>  
A;Cross-references: EMBL:X14053  
R;Perlmutter, R.M.; Marth, J.D.; Lewis, D.B.; Peet, R.; Ziegler, S.F.; Wilson, C.B.  
J. Cell. Biochem. 38, 117-126, 1988  
A;Title: Structure and expression of lck transcripts in human lymphoid cells.  
A;Reference number: S07822; MUID:89123626; PMID:3265417  
A;Accession: S07822

A;Molecule type: mRNA  
A;Residues: 1-86, 'P', 88-509 <PER>  
A;Cross-references: EMBL:X13529; NID:G34294; PIDN:CAA31884.1; PID:G34295  
R;Koga, Y.; Caccia, N.; Toyonaga, B.; Spolski, R.; Yanagi, Y.; Yoshikai, Y.; Mak, T.W.  
Eur. J. Immunol. 16, 1643-1646, 1986  
A;Title: A human T cell-specific cDNA clone (YT16) encodes a protein with extensive homol  
A;Reference number: S07200; MUID:87133831; PMID:3493153  
A;Accession: S07200  
A;Molecule type: mRNA  
A;Residues: 1-205, 'ASAITPI', 212-257, 'RCGW', 262, 'TTT', 266, 'T', 268-281, 'AGRIPL', 287-503, 'ST'  
A;Cross-references: EMBL:X05027; NID:G36807; PIDN:CAA28691.1; PID:G36808  
R;Veillette, A.; Foss, F.M.; Sausville, E.A.; Bolen, J.B.; Rosen, N.  
Oncogene Res. 1, 357-374, 1987  
A;Title: Expression of the lck tyrosine kinase gene in human colon carcinoma and other n  
A;Reference number: S01879; MUID:88217332; PMID:2835736  
A;Accession: S01879  
A;Molecule type: mRNA  
A;Residues: 368-471, 'H', 473-509 <VEI>  
A;Cross-references: EMBL:X06369; NID:G34288; PIDN:CAA29667.1; PID:G34289  
R;Trevisan, J.M.; Lin, Y.; Chen, S.J.; Phillips, C.A.; Canna, C.; Linna, T.J.  
Biochim. Biophys. Acta 888, 286-295, 1986  
A;Title: Human T lymphocytes express a protein-tyrosine kinase homologous to p56 (LSTRA).  
A;Reference number: S07143; MUID:87000726; PMID:3489436  
A;Accession: S07143  
A;Molecule type: mRNA  
A;Residues: 'A', 376-509 <TRE>  
A;Cross-references: EMBL:X04476; NID:G35779; PIDN:CAA28165.1; PID:G35780  
R;Takadera, T.; Leung, S.; Gernone, A.; Koga, Y.; Takiyama, Y.; Miyamoto, N.G.; Mak, T.W.  
Mol. Cell. Biol. 9, 2173-2180, 1989  
A;Title: Structure of the murine lck gene and its rearrangement in a murine lymphoma cell  
A;Reference number: A32797; MUID:89313764; PMID:2787474  
A;Accession: A32797  
A;Molecule type: DNA  
A;Residues: 1-35 <TAK>  
A;Cross-references: GB:M26692; NID:G341523; PIDN:AAA59503.1; PID:G349702  
R;Garvin, A.M.; Pawar, S.; Marth, J.D.; Perlmutter, R.M.  
Mol. Cell. Biol. 8, 3058-3064, 1988  
A;Title: Structure of the murine lck gene and its rearrangement in a murine lymphoma cell  
A;Reference number: I57636; MUID:89096891; PMID:2850479  
A;Accession: I57636  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-35, 'VR', <RES>  
A;Cross-references: GB:M21510; NID:G187031; PIDN:AAA9501.1; PID:G553522  
C;Comment: Protein tyrosine kinases play important roles in the control of cell growth ar  
C;Genetics:  
A;Gene: GDB:LCK  
A;Cross-references: GDB:119360; OMIM:153390  
A;Map position: 1p35-1p34.3  
A;Introns: 35/3; 63/1; 93/2; 126/2; 161/1; 211/1; 262/1; 322/1; 347/3; 399/1; 443/1  
C;Function:  
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 ho  
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho  
F;2-509/Product: protein-tyrosine kinase lck #status predicted <MAT>  
F;68-116/Domain: SH3 homology <SH3>  
F;127-224/Domain: SH2 homology <SH2>  
F;243-501/Domain: protein kinase homology <KIN>  
F;251-259/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;3.5/Binding site: palmitate (Cys) (covalent) #status predicted  
F;273/Active site: Lys #status predicted  
F;394,505/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

Query Match 25.1%; Score 340; DB 1; Length 509;  
Best Local Similarity 40.6%; Pred. No. 2.3e-20;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

QY 26 VTMEERSKAT-----AVALGSPFAGGPAELSLRLGEPLTIVSEDDGWTVLSEVSGRE 79  
Db 49 VTTEGSPFPASPQDNLVIALHSYFSDHGDGLCFEKGQRLILEQGEKWKQASLTIGQE 108  
QY 80 YNIPSVHVGKVS-----HGWLYEGLSRKAEELLLLPNPGGAFIRSQTRRSYSLSVR 135

109	GFIPFNVAKNSLEPEPWFKNLSRDKAERQLLAPGNTHGSEFLIRESESTAGSFLSVR	168
Db		
136	LSRPASWDRIHRYTHCLDNGWLIISRLTFPSLQALVDHYSLADDICLLKEPCTVLR	195
Qy		
169	DPDQNGGEVWHYKIRNLNDGGFVISRIPTFPGHLELVHVTNASGLCTLSRRCQOK	228
Db		

RESULT 11

I48845  
protein-tyrosine kinase (EC 2.7.1.112) lck, lymphocyte - mouse  
N/Alternate names: p56; protein-tyrosine kinase tck  
C/Species: Mus musculus (house mouse)  
C/Date: 18-Feb-2000 #sequence revision 18-Feb-2000 #text\_change 03-Mar-2000  
C/Accession: I48845; A23639; I57629; I77452  
R/Voronova, A.F.; Sefton, B.M.  
Nature 319, 682-685, 1986  
A/Title: Expression of a new tyrosine protein kinase is stimulated by retrovirus promote  
A/Reference number: I48845; MUID:86146842; PMID:3081813  
A/Accession: I48845  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-509 <VORI>  
A/Cross-references: EMBL:X03533; NID:g54813; PIDN:CAA27234.1; PID:g54814  
R/Marth, J.D.; Peet, R.; Krebs, E.G.; Perlmutter, R.M.  
Cell 43, 393-404, 1985  
A/Title: A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpress  
A/Reference number: A23639; MUID:86079521; PMID:2416464  
A/Accession: A23639  
A/Molecule type: mRNA  
A/Residues: 1-282, 'VP', 285-509 <NAR>  
A/Cross-references: GB:M12056; NID:gl98763  
A/Note: the sequence is revised in GenBank entry MUSLCK, release 116.0, (PIDN:AAB59674.1  
R/Voronova, A.F.; Adler, H.T.; Sefton, B.M.  
Mol. Cell. Biol. 7, 4407-4413, 1987  
A/Title: Two lck transcripts containing different 5' untranslated regions are present in  
A/Reference number: I57629; MUID:88142832; PMID:3501824  
A/Accession: I57629  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-11 <VOR>  
A/Cross-references: GB:M18098; NID:gl98766; PIDN:AAA39421.1; PID:gl98767  
R/Garvin, A.M.; Pawar, S.; Marth, J.D.; Perlmutter, R.M.  
Mol. Cell. Biol. 8, 3058-3064, 1988  
A/Title: Structure of the murine lck gene and its rearrangement in a murine lymphoma cell  
A/Reference number: I57636; MUID:89096891; PMID:2850479  
A/Accession: I77452  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-35, VR <GAR>  
A/Cross-references: GB:M21511; NID:gl98768; PIDN:AAA39422.1; PID:g554186  
C/Speciesfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h  
C/Keywords: APP; autophosphorylation; blocked amino end; kinase-related transforming pro  
F/68-116/Domain: SH3 homology <SH3>  
F/127-224/Domain: SH2 homology <SH2>  
F/243-501/Domain: protein kinase homology <KIN>  
F/251-359/Region: protein kinase Atp-binding motif  
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F/273/Active site: -ys #status predicted  
F/394-505/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 25.0%; Score 338.5; DB 1; Length 509;  
Best Local Similarity 39.1%; Pred. No. 3e-20;  
Matches 77; Conservative 28; Mismatches 85; Indels 7; Gaps 2

Qy	3	S	I	P	S	R	K	S	L	P	S	S	S	V	Q	G	P	V	T	M	E	R	S	K	A	T	A	V	A	L	G	S	F	P	A	G	G	A	E	S	L	S	L	G	E	P	L	T	V	62					
Db	35	S	I	P	I	R	G	S	E	V	R	D	L	-	-	-	V	T	E	G	S	P	P	A	S	P	L	O	D	N	I	V	I	A	L	H	S	E	P	S	H	D	G	L	G	F	E	K	E	G	E	L	R	L	91
Qy	63	S	E	D	G	M	W	T	V	L	S	V	S	G	R	E	N	I	S	V	H	V	K	S	-	-	-	H	G	M	L	E	G	L	S	R	E	K	A	E	L	L	L	P	N	G	E	G	A	F	118				
Db	92	E	O	S	G	E	M	W	A	K	S	L	T	T	G	E	G	F	I	P	N	F	V	A	K	A	N	S	L	E	P	E	P	W	F	K	N	L	S	K	D	A	E	R	O	L	L	A	P	N	T	H	S	F	151

[illegible]

## RESULT 12

A39939 protein-tyrosine kinase (EC 2.7.1.112) tkl [similarity] - chicken  
N;Alternate names: kinase-related transforming protein (tkl); T-cell surface antigen associated protein (TSP)  
C;Species: Gallus gallus (chicken)  
C;Date: 16-Jun-2000 #sequence revision 16-Jun-2000 #text\_change 16-Jun-2000  
C;Accession: A42126; A39939  
R;Chow, L.M.; Ratcliffe, M.J.; Veillette, A.  
Mol. Cell. Biol. 12, 1226-1233, 1992  
A;Title: tkl is the avian homolog of the mammalian lck tyrosine protein kinase gene.  
A;Reference number: A42126; MUID:92186854; PMID:1545804  
A;Accession: A42126  
A;Molecule type: mRNA  
A;Residues: 1-88 <CHO>  
A;Cross-references: GB:M85043  
A;Experimental source: thymus, spleen  
A;Note: sequence extracted from NCBI backbone (NCBIN:89831, NCBIPI:89833)  
R;Streibhardt, K.; Mullins, J.I.; Bruck, C.; Ruebeamen-Waigmann, H.  
Proc. Natl. Acad. Sci. U.S.A. 84, 8778-8782, 1987  
A;Title: Additional member of the protein-tyrosine kinase family: the src-and lck-related protein-tyrosine kinase  
A;Reference number: A39939; MUID:8097370; PMID:3321053  
A;Accession: A39939  
A;Molecule type: mRNA  
A;Residues: 52-507 <STR>  
A;Cross-references: GB:J03579; NID:G212712; PID:AAA49081.1; PID:G212713  
C;Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 homology  
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phosphorylation; protein-tyrosine kinase  
F;66-114/Domain: SH3 homology <SH3>  
F;125-222/Domain: SH2 homology <SH2>  
F;241-499/Domain: protein kinase ATP-binding motif  
F;249-257/Region: protein kinase ATP-binding motif  
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
F;392/503/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 24.6%; Score 333; DB 1; Length 507;  
Best Local Similarity 43.2%; Pred. No. 8.7e-20;  
Matches 70; Conservative 27; Mismatches 61; Indels 4; Gaps 1;

Qy	38	VALGSPAGGPAELSRULGEPLTIVSDGDWTVTVLSEVSGRENYNIISVTHVKYS-----HG	93
Db	65	VAILYDEPTHDGLGKQGEKURLVSESGWMPAQSLTTQCGELIHNFMVWNSLEPEP	124
Qy	94	WLYEGISREKAEELLLLPNPGGAFLIRSESQTRGSYSVLUSRSPASWDRIHRYIHCL	153
Db	125	WFFKNLSRKNAAEARLLASNGTHGSFTRIRESYTSKGSYSVSRVDFDQNGQGVTHYKIRNM	184
Qy	154	DNQWLVIISRLTPPSLQALVDHYSELADDIICLLKEPVCVLQR	195
Db	185	DNGGYIISPRVTSSLHELVEVYSSSDGLCTRLGKPCRTQK	226

RESULT 13

protein-tyrosine kinase (BC 2.7.1.112) fyk - Pacific electric ray  
B49114  
C:Species: Torpedo californica (Pacific electric ray)  
C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 18-Feb-2000  
C:Accession: B49114  
R:Swope, S.L.; Hugarir, R.L.  
J. Bio. Chem. 258, 25152-25161, 1993  
A:Title: Molecular cloning of two abundant protein tyrosine kinases in Torpedo electric ray  
A:Reference number: A49114; MUID:94043386; PMID:8227079  
A:Accession: B49114  
A:Status: preliminary  
A:Molecule type: mRNA



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 08:16:06, Search time 39 Seconds  
(without alignments)  
348.469 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSLSRRKSPSPSLSSV.....RESLSFYISUNDEAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1342	99.2	260	1 SLA2_HUMAN	Qh6q3 homo sapien
2	1023	75.6	258	1 SLA2_MOUSE	Q8r410 mus musculus
3	501	37.0	312	1 SLA1_RAT	P59622 rattus norv
4	491.5	36.3	280	1 SLA1_MOUSE	Q60898 mus musculus
5	481.5	35.6	275	1 SLA1_HUMAN	Q13239 homo sapien
6	370.5	27.4	511	1 LYN_HUMAN	P07948 homo sapien
7	367.5	27.2	511	1 LYN_MOUSE	P25911 mus musculus
8	367.5	27.2	511	1 LYN_RAT	Q07014 rattus norv
9	360.5	26.6	525	1 HCK_HUMAN	P08631 homo sapien
10	358	26.5	503	1 HCK_MACFA	Q5m30 macaca fasc
11	352.5	26.1	504	1 BLK_HUMAN	Q5m30 homo sapien
12	352	26.0	523	1 HCK_MOUSE	P08103 mus musculus
13	351.5	26.0	502	1 HCK_RAT	P50545 rattus norv
14	341.5	25.2	498	1 BLK_MOUSE	P16277 mus musculus
15	340	25.1	508	1 LCK_HUMAN	P06239 homo sapien
16	338.5	25.0	508	1 LCK_MOUSE	P06240 mus musculus
17	333	24.6	507	1 LCK_CHICK	P22683 gallus gall
18	332	24.5	517	1 SR42_DROME	Q9V9J3 grosophila
19	319.5	23.6	537	1 YES_XENLA	P10936 xenopus lae
20	317	23.4	541	1 YES_CHICK	P09324 gallus gall
21	315.5	23.3	536	1 FYN_HUMAN	P06241 homo sapien
22	315.5	23.3	543	1 YES_HUMAN	P07947 homo sapien
23	312.5	23.1	529	1 FGR_HUMAN	P09769 homo sapien
24	311	23.0	544	1 YES_XPHE	P27447 xiphophorus
25	310.5	22.9	536	1 FYN_XENLA	P13405 xenopus lae
26	309	22.8	528	1 YES_AVISR	P00527 avian sarco
27	309	22.8	533	1 FYN_MOUSE	P39688 mus musculus
28	306.5	22.7	541	1 YES_MOUSE	Q04736 mus musculus
29	305.5	22.6	536	1 FYN_XPHE	P27446 xiphophorus
30	303	22.4	517	1 FGR_MOUSE	P14234 mus musculus
31	301	22.2	535	1 VRK_CHICK	Q02977 gallus gall
32	300	22.2	539	1 YES_CANFA	Q28923 canis fami
33	297.5	22.0	533	1 FYN_CHICK	Q05876 gallus gall

34	290	21.4	506	1 SRK4_SPOLA	P42890 sporogilla 1
35	286.5	21.2	526	1 SRC_AVISR	P00525 avian sarco
36	286.5	21.2	532	1 SRC_CHICK	P00523 gallus gall
37	286.5	21.2	552	1 SR64_DROME	P00528 drosophila
38	286.5	21.2	557	1 SRC_AVISR	P14085 avian sarco
39	286.5	21.2	587	1 SRC_AVIS2	P15054 avian sarco
40	284.5	21.0	568	1 SRC_AVIS3	P14084 avian sarco
41	277.5	20.5	535	1 SRC_RAT	Q9W0D9 rattus norv
42	276.5	20.4	526	1 SRC_RSVSR	P00524 rous sarcom
43	276.5	20.4	535	1 SRC_HUMAN	P12931 homo sapien
44	276	20.4	509	1 STK_HYDRA	P17713 hydra atten
45	274.5	20.3	531	1 SRC2_XENLA	P13116 xenopus lae

## ALIGNMENTS

RESULT 1					
SLA2_HUMAN					
ID	SLA2_HUMAN	STANDARD;	PRT;	260	AA.
AC	Q9H6Q3; Q96Y18; Q96Q14; Q9H135;				
DT	10-OCT-2003 (Rel. 42, Created)				
DT	10-OCT-2003 (Rel. 42, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Src-like-adaptor 2 (Src-like adapter protein-2) (SLAP-2) (Modulator of antigen receptor signaling) (MARS).				
GN	SLA2 OR SLAP2 OR C20CRF156.				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1), CHARACTERIZATION, FUNCTION MYRISTOYLATION, INTERACTION WITH CBL, AND MUTAGENESIS OF GLY-1.				
RX	MEDLINE=21553259; PubMed=11696592;				
RA	Holland S.J., Liao X.C., Mendenhall M.K., Zhou X., Pardo J., Chu P., Spencer C., Fu A.C., Sheng N., Yu P., Pali E., Nagin A., Shen M., Yu S., Chan E., Wu X., Li C., Wolsettschlag M., Aversa G., Kolbinger F., Bennett M.K., Molineaux S., Luo Y., Payan D.G., Mancebo H.S.Y., Wu J.,				
RT	"Functional cloning of Src-like adapter protein-2 (SLAP-2), a novel inhibitor of antigen receptor signaling."				
RL	J. Exp. Med. 194:1263-1276 (2001).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND ALTERNATIVE INITIATION.				
RC	TISSUE=Thymus;				
RX	MEDLINE=22415750; PubMed=12527895;				
RA	Lorsto M.P., McGlade C.J.,				
RT	"Cloning and characterization of human Src-like adaptor protein 2 and a novel splice isoform, SLAP-2-v."				
RL	Oncogene 22:266-273 (2003).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Hepatoma;				
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Ora T., Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;				
RT	"NEDD human cDNA sequencing project."				
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21638749; PubMed=11780052;				
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L., Bailey J., Barlow K.P., Bates K.N., Bead L.M., Beare D.M., Beasley O.P., Bird C.P., Blake S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,				

RA Hammond S., Hazley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehtvasilho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McComachie L.J., McDay K., McMurray A.,  
RA Milne S.A., Mastry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
RA Skuce C.D., Smith M.I., Soderlund C., Steward C.A., Sulton J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Walli M., Wallis J.M., Williams S.A.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate; PubMed=12477932;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Eustow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Locquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RX CHARACTERIZATION.  
RP MEDLINE=22013997; PubMed=11891219;  
RA Pandey A., Ibarola N., Kratchmarova I., Fernandez M.M.,  
RA Constantinescu S.N., Ohara O., Sawasdikosol S., Lodish H.F., Mann M.,  
RT "A novel Src homology 2 domain-containing molecule, Src-like adapter  
RT protein-2 (SLAP-2), which negatively regulates T cell receptor  
RT signaling.";  
RL J. Biol. Chem. 277:19131-19138(2002).  
CC -!- FUNCTION: Adapter protein, which negatively regulates T-cell  
CC receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced  
CC activation of nuclear factor of activated T-cells. May act by  
CC linking signaling proteins such as ZAP70 with CBL, leading to a  
CC CBL dependent degradation of signaling proteins.  
CC -!- SUBUNIT: Interacts with phosphorylated proteins ZAP70 and CD32 via  
CC its SH2 domain (By similarity). Interacts with phosphorylated CBL  
CC via its C-terminal domain.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Isoform 1 is localized to the  
CC plasma membrane and intracellular vesicles, including late  
CC endosomal vesicles. Isoform 2 may be cytoplasmic and is not  
CC localized to membranes.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1; Synonyms=sp28;  
CC IsoId=Q9H6Q3-1; Sequence=Displayed;  
CC Note=Isoform 3 is produced by alternative initiation at  
CC Met-27 of isoform 1;  
CC Name=2; Synonyms=sp23, SLAP-2-V, MAR8-V;  
CC IsoId=Q9H6Q3-2; Sequence=VSP\_007240, VSP\_007241;  
CC Note=Isoform 4 is produced by alternative initiation at  
CC Met-27 of isoform 2;  
CC Event=Alternative initiation;  
CC

CC Comment=4 isoforms, 1/p28 (shown here), 2/p23, 3/p25 and 4/p20,  
CC ARE produced by alternative initiation at Met-0 and Met-27;  
CC -!- TISSUE SPECIFICITY: Predominantly expressed in immune system, with  
CC highest levels in peripheral blood leukocytes. Expressed in  
CC spleen, thymus and lymph nodes. Expressed in T-cells as well as in  
CC monocytes, and at low level in B-cells. Also detected in placenta,  
CC prostate, skin, retina and colon.  
CC -!- DOMAIN: The loss of the C-terminal domain partially abolishes the  
CC inhibitory function, but can be partially compensated by higher  
CC level of protein expression.  
CC -!- PTM: May be phosphorylated (By similarity).  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
CC  
CC EMBL; AF26353; AAL29204.1; -;  
CC EMBL; AF290985; AAL38197.1; -;  
CC EMBL; AF290986; AAL38198.1; -;  
CC EMBL; AK025645; BAB15201.1; -;  
CC EMBL; AL031662; CAC44645.1; ALT\_TERM.  
CC EMBL; AL050318; CAB75365.1; -;  
CC EMBL; BC042041; AAH42041.1; -;  
CC HSSP; P06239; ILKK.  
CC Genew; HGNC:17329; SLA2.  
CC MIM; 606577; -;  
CC InterPro; IPR000980; SH2.  
CC InterPro; IPR001452; SH3.  
CC Pfam; PF00017; SH2; 1.  
CC Pfam; PF00018; SH3; 1.  
CC PRINTS; PR0401; SH2DOMAIN.  
CC SMART; SM00093; SH2; 1.  
CC SMART; SM00252; SH2; 1.  
CC SMART; SM00326; SH3; 1.  
CC PROSITE; PS00001; SH2; 1.  
CC PROSITE; PS00002; SH3; 1.  
CC Membrane; SH2 domain; SH3 domain; Myristate; Phosphorylation;  
CC Alternative initiation; Alternative splicing; Lipoprotein.  
CC BY SIMILARITY.  
CC INIT MET 0 0  
CC CHAIN 1 260  
CC SRC-LIKE-ADAPTER 2, ISOFORM 1 AND ISOFORM  
CC 2.  
CC SRC-LIKE-ADAPTER 2, ISOFORM 3 AND ISOFORM  
CC 4.  
CC FOR ISOFORM 3 AND ISOFORM 4.  
CC SH3.  
CC SH2.  
CC SLA C-TERMINAL.  
CC N-mristoyl glycine (By similarity).  
CC LADICLLKEPCVQLQAGPLGKQIPLVTV -> GWPAP  
CC WQGYTTCDAEDTTQLERAGQLPPVF (in isoform  
CC 2).  
CC /FTid-VSP\_007240.  
CC /FTid-VSP\_007241.  
CC /FTid-VSP\_007241.  
CC G->A: ABOLISHES LOCALIZATION TO  
CC MEMBRANES.  
CC SEQUENCE 260 AA; 28454 MW; A402C03449261B3D CRC64;  
CC  
CC Query Match 99.2%; Score 1342; DB 1; Length 260;  
CC Best Local Similarity 99.6%; Pred. No. 3.7e-105;  
CC Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
CC  
CC QY 2 GSLPSRKSLPSLSLSSVQGVPTWEAERSKATAVALGSPAGGPAELSLRLSEPLTI 61  
CC DB 1 GSLPSRKSLPSLSLSSVQGVPTWEAERSKATAVALGSPAGGPAELSLRLSEPLTI 60  
CC QY 62 VSEDGDWMTVLSEVSGREYNIPSHVGVKSHGWLVEGLSREKAEILLLPNGPGGAFTIR 121

Db 61 VSEDGDMWTVLSEVSREYNISVHVAKVSHGWLVEGLSREKAEILLPLGPGGAFLIR 120  
 QY 122 ESOTRGYSLSVRLSRPASWDRIRHYRHCLDNGWLYISPLTPPSIQALVDHYSELAD 181  
 Db 121 ESOTRGYSLSVRLSRPASWDRIRHYRHCLDNGWLYISPLTPPSIQALVDHYSELAD 180  
 QY 182 DICCLKEPCVLQACGLPGKPIPLVTVQRTPLNWKELDSLLSEATGSGESLLSEGL 241  
 Db 181 DICCLKEPCVLQACGLPGKPIPLVTVQRTPLNWKELDSLLSEATGSGESLLSEGL 240  
 QY 242 RESLSFYISLNDVAVSLDDA 261  
 Db 241 RESLSFYISLNDVAVSLDDA 260

RESULT 2  
 SLA2\_MOUSE STANDARD; PRT; 258 AA.  
 AC Q8R410; Q8COK2; Q8V142; Q9D1Z9;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Src-like-adaptor 2 (Src-like adaptor protein-2) (SLAP-2).  
 GN SLA2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), ALTERNATIVE INITIATION,  
 RP CHARACTERIZATION, FUNCTION, MYRISTOYLATION, PHOSPHORYLATION,  
 RP INTERACTION WITH ZAP70 AND CBL, AND MUTAGENESIS OF GLY-1; MET-26 AND  
 RP ARG-119.  
 RX MEDLINE=22022020; PubMed=12024036;  
 RA Loreto M.P., Berry D.M., McGlade C.J.;  
 RT "Functional cooperation between c-Chl and Src-like adaptor protein 2  
 RT in the negative regulation of T-cell receptor signaling.";  
 RL Mol. Cell. Biol. 22:4241-4255(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, MYRISTOYLATION,  
 RP INTERACTION WITH CBL; ZAP70 AND CD3Z, AND MUTAGENESIS OF GLY-1; PRO-81  
 RP ARG-119.  
 RX MEDLINE=22013997; PubMed=11891219;  
 RA Pandey A., Ibarrola N., Kratchmarova I., Fernandez M.M.,  
 RA Constantinescu S.N., Ohara C., Sawadkosol S., Lodish H.F., Mann M.;  
 RT "A novel Src homology 2 domain-containing molecule, Src-like adaptor  
 RT protein-2 (SLAP-2), which negatively regulates T cell receptor  
 RT signaling.";  
 RL J. Biol. Chem. 277:19131-19138(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=C57BL/6J, and NOD; TISSUE=Retina, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmer S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltas L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., King B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Vervardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 CC -!- FUNCTION: Adapter protein, which negatively regulates T-cell  
 CC receptor (TCR) signaling. Inhibits T-cell antigen-receptor induced  
 CC activation of nuclear factor of activated T-cells. May act by  
 CC linking signaling proteins such as ZAP70 with CBL, leading to a  
 CC CBL dependent degradation of signaling proteins.  
 CC -!- SUBUNIT: Interacts with phosphorylated CBL via its C-terminal  
 CC domain. Interacts with phosphorylated proteins ZAP70 and CD3Z via  
 CC its SH2 domain.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; localized to the plasma  
 CC membrane and intracellular vesicles, including late endosomal  
 CC vesicles.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative initiation;  
 CC Comment=2 isoforms, 1/p28 (shown here) and 2/p25, are produced  
 CC by alternative initiation at Met-0 and Met-26;  
 CC -!- TISSUE SPECIFICITY: Mainly expressed in immune system. Highly  
 CC expressed in spleen and thymus and expressed at intermediate  
 CC levels in lung. Not expressed in liver, heart and brain. Isoform 1  
 CC is predominant in lung and spleen, while isoform 2 is predominant  
 CC in thymus.  
 CC -!- DOMAIN: The loss of the C-terminal domain partially abolishes the  
 CC inhibitory function.  
 CC -!- PTM: May be phosphorylated.  
 CC -!- MISCELLANEOUS: Ref.1 confirmed the alternative initiation by  
 CC mutating the Met in position -1 to Val, and showed that isoform 1  
 CC is abolished in favor of isoform 2.  
 CC -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF287467; AAL38196.1; -;  
 CC EMBL; AF434990; AAL86403.1; -;  
 CC EMBL; AK020837; BAB32223.1; AUT\_INIT.  
 CC EMBL; AK030877; BAC37168.1; -;  
 CC EMBL; AK088672; BAC40495.1; -;  
 CC HSSP; P06239; 1LKK.  
 CC MGD; MGI:1325049; Sla2.  
 CC GO; GO:0015023; Cytoplasmic vesicle; IDA.  
 CC GO; GO:0005770; Late endosome; IDA.  
 CC GO; GO:0005886; C:plasma membrane; IDA.  
 CC GO; GO:0005515; F:protein binding; IDA.  
 CC GO; GO:0042110; F:T-cell activation; IDA.  
 CC InterPro; IPR000980; SH2.  
 CC InterPro; IPR001452; SH3.  
 CC Pfam; PF00017; SH2; 1.  
 CC Pfam; PF00018; SH3; 1.  
 CC PRINTS; PR00401; SH2DOMAIN.  
 CC ProDom; PD000093; SH2; 1.  
 CC SMART; SM00252; SH2; 1.  
 CC SMART; SM00326; SH3; 1.  
 CC PROSITE; PS50001; SH2; 1.  
 CC PROSITE; PS50002; SH3; 1.  
 CC Immune response; Membrane; SH2 domain; SH3 domain; Myristate;  
 KW Phosphorylation; Alternative initiation; Lipoprotein.  
 FT INIT MET 0  
 FT CHAIN 1 258 SRC-LIKE-ADAPTER 2, ISOFORM 1.

```

FT CHAIN 26 258 SRC-LIKE-ADAPTER 2, ISOFORM 2.
FT INIT MET 26 26 FOR ISOFORM 2.
FT DOMAIN 30 90 SH3.
FT DOMAIN 92 189 SH2.
FT DOMAIN 189 258 SLA C-TERMINAL.
FT LIPID 1 1 N-myristoyl glycine.
FT MUTAGEN 1 1 G->A: ABOLISHES LOCALIZATION TO MEMBRANES.
FT MUTAGEN 26 26 M->V: ABOLISHES ISOFORM 2.
FT MUTAGEN 81 81 P->L: DOES NOT AFFECT ITS INHIBITORY FUNCTION.
FT MUTAGEN 119 119 R->E: ABOLISHES INTERACTION WITH ZAP70, AND ITS INHIBITORY FUNCTION.
FT CONFLICT 127 127 C -> Y (IN REF. 3; BAC27168).
FT CONFLICT 159 159 S -> T (IN REF. 3; BAC27168).
FT CONFLICT 160 160 P -> H (IN REF. 2).
SQ SEQUENCE 258 AA; 28345 MW; 8270E93DE3FC696A CRC64;

Query Match 75.6%; Score 1023; DB 1; Length 258;
Best Local Similarity 79.3%; Pred. No. 1.8e-78;
Matches 207; Conservative 16; Mismatches 34; Indels 4; Gaps 3;

QY 2 GSLSRKRKSLPSLSASSVQGGPVTMEARSKATAVALGSPAGGPAELSLRLGELPLTI 61
Db 1 GSLSRKRKTSFSPSPSSSGPDPVSPQPERHKVTAVALGSPAGGPAELSLRLGELPLTI 59

QY 62 VSEGDGDMWTVLSEVSGREYNIPSHVGVKSHGWLVEGLGREKAEELLPLGNPGGAFLIR 121
Db 60 ISEGDGDMWTVOSEVSGREYHMPVYVAKVAHGWLYEGLGREKAEELLPLGNPGGAFLIR 119

QY 122 ESQTRRSYSLVRLSPASWDRIHRYIHCLDNGWLISPRITFPSLOALVDHYSELAD 181
Db 120 ESQTRRCYSYSLVRLSPASWDRIHRYIQRDLONGWLISPRITFPSLHALVEHSELAD 179

QY 182 DICLLKEPCVLQAGLPKQIPPLVTVQRTPLNWKELDSSLFSEA-ATGEESLLSFG 240
Db 180 GICCPLEPCVLQGLPGKQDPPTVPTSSLNKKLDRSLFLFLFLEAPASGEASLLSFG 239

QY 241 LRESLSFYISLNDFAVSLDDA 261
Db 240 LRESLSFYISLAED--PLDDA 258

RESULT 3
SL1-RAT SL1-RAT STANDARD; PRT; 312 AA.
AC PS3622;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SRC-like-adaptor (Src-like-adaptor protein 1).
GN SLA OR SLAP OR SLAP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Afragias P., Molnar Z., Parnavelas J.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adaptor protein, which negatively regulates T-cell
CC activation of nuclear factor of activated T-cells. Involved in the
CC negative regulation of positive selection and mitosis of T-cells.
CC May act by linking signaling proteins such as ZAP70 with CBL,
CC leading to a CBL dependent degradation of signaling proteins (By
CC similarity).
CC -!- SUBUNIT: Homodimer. Interacts with phosphorylated CBL, SYK and
CC LAT. Homodimerization and interaction with phosphorylated CBL
CC occurs via its C-terminal domain. Interacts with
CC PDGFRB and EphA2. Interacts with phosphorylated proteins ZAP70;
CC CD32; VAV1 and LCP2 via its SH2 domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; colocalizes with endosomes (By

```

```

similarity).
CC -!- DOMAIN: The C-terminal domain is essential for the
CC homodimerization and the interaction with CBL. While the
CC interaction with CBL is apparently mediated via the hydrophobic
CC region of this domain, the highly charged region is apparently
CC required for the homodimerization (By similarity).
CC -!- PTM: Phosphorylated (By similarity).
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AY217759; AAC61134.1; -
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001452; SH3.
CC Pfam; PF00017; SH2; 1.
CC PRINTS; PR00018; SH3; 1.
CC PRODOM; PD000093; SH2; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS50001; SH2; 1.
CC PROSITE; PS50002; SH3; 1.
CC SH2 domain; SH3 domain; Phosphorylation.
KW DOMAIN 38 98 SH3.
FT DOMAIN 100 191 SH2.
FT DOMAIN 206 312 SLA C-TERMINAL.
SQ SEQUENCE 312 AA; 34744 MW; A05412D39AE68388 CRC64;

Query Match 37.0%; Score 501; DB 1; Length 312;
Best Local Similarity 41.3%; Pred. No. 1.1e-34;
Matches 112; Conservative 41; Mismatches 90; Indels 28; Gaps 6;

QY 11 LPSPSLSSVQGG-----PVTMEARSKATAV-----ALGSFPAGGPAELSLRLGEP 58
Db 5 LFGPSTSRGKEMGNMKSSTAPLERPLNTEGLSDFLAVLNDYPSDIPSPFRGEX 64

QY 59 LTVSEGDGDMWTVLSEVSGREYNIPSHVGVKSHGWLVEGLGREKAEELLPLGNPGGAFLIR 118
Db 65 LRVISDEGGWKAISLSLTGRESYIPGICVARVYHGWLFEGGLGRDKABELLQPLTKTGSF 124

QY 119 LIRESDTRSGYSLVRLSPASWDRIHRYIHCLDNGWLISPRITFPSLOALVDHYSE 178
Db 125 MIRESETKGPFYSLVR-----HRQVXHYIFRLPNNWYIISPRITFQCLEDLVTHYSE 178

QY 179 LADDICLLKEPCVLQR-----AGLPKQIPPLVTVQRTPLNWKELDSSLFSEAATG- 232
Db 179 VADGLCCVLTTPCLAQNTAPTAQPSCTSPGSPVTLQKTFDKRVSSLOEGPEGAENP 238

QY 233 ----EESLSEGLRESLSFYISL-NDEAVSLD 259
Db 239 LRVDLSFYGLRESIASYLSLTGDDSNFD 269

RESULT 4
SL1-MOUSE SL1-MOUSE STANDARD; PRT; 280 AA.
AC Q60838; Q8C908; Q8CAT0; Q8CBE9; Q8QZX8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SRC-like-adaptor (Src-like-adaptor protein 1) (mslap).
GN SLA OR SLAP OR SLAP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

[1] SEQUENCE FROM N.A. (ISOPFORM 1), TISSUE SPECIFICITY, AND INTERACTION  
 RP WITH EPHA2.  
 RP TISSUE=Embryonic brain;  
 RX MEDLINE=95370243; PubMed=7543898;  
 RA Pandey A., Duan H., Dixit V.M.;  
 RT "Characterization of a novel Src-like adapter protein that associates  
 RL with the Eok receptor tyrosine kinase";  
 RL J. Biol. Chem. 270:19201-19204 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOPFORM 1).  
 RX MEDLINE=20093395; PubMed=10630289;  
 RA Carrier A., Nguyen C., Victorero G., Granjeaud S., Rocha D.,  
 RA Bernard K., Mlizek A., Ferrier P., Mallissen M., Naquet P.,  
 RA Mallissen B., Jordan B.R.;  
 RT "Differential Gene expression in CD3epsilon- and RAG1-deficient  
 RL thymuses: definition of a set of genes potentially involved in  
 RL thymocyte maturation";  
 RL Immunogenetics 50:255-270 (1999).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOPFORM 1).  
 RX STRAIN=ILS, and ISS;  
 RX MEDLINE=21363810; PubMed=11471062;  
 RA Ehringer M.A., Thompson J., Conroy C., Xu Y., Yang F., Canniff J.,  
 RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;  
 RT "High-throughput sequence identification of gene coding variants  
 RL within alcohol-related QTLs";  
 RL Mamm. Genome 12:657-663 (2001).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOPFORMS 1 AND 2).  
 RX STRAIN=C57BL/6J; TISSUE=Cerebellum, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Kikado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Kasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Balgarelli R., Hill D.P., Sult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matcud H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.B., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedziński R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovskiy N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempé C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Varadar R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynnshaw-Boris A., Yanagisawa M., Yang L., Yang L.,  
 RA Yuan Z., Zavalova M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of c  
 RL 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOPFORM 1).  
 RX TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altshul S.F., Zengerb B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haich F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Teshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Lomuelano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

CC required for the homodimerization (By similarity).

CC -1- PTH: Phosphorylated (By similarity).

CC -1- MISCELLANEOUS: SLA deficient mice show a strong upregulation

CC of TCR and CD5 at the CD4(+) CD8(-) stage, and an enhanced

CC positive selection in T-cells.

CC -1- SIMILARITY: Contains 1 SH2 domain.

CC -1- SIMILARITY: Contains 1 SH3 domain.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U29056; AAA82756.1; -

DR EMBL; AJ131777; CAB66139.1; -

DR EMBL; AY079449; AAL87537.1; -

DR EMBL; AY079450; AAL87538.1; -

DR EMBL; AK036167; BAC29328.1; -

DR EMBL; AK037901; BAC29896.1; -

DR EMBL; AK041565; BAC30988.1; -

DR EMBL; BC032922; AAH32922.1; -

DR HSP; P18277; IBLK.

DR MGD; MG1104295; Sla.

DR InterPro; IPR000980; SH2.

DR InterPro; IPR001452; SH3.

DR Pfam; PF00017; SH2; 1.

DR Pfam; PF00018; SH3; 1.

DR PRINTS; PR00401; SH2DOMAIN.

DR PRODOM; PD000093; SH2; 1.

DR SMART; SM00326; SH3; 1.

DR SMART; SM00326; SH3; 1.

DR PROSITE; PS00001; SH2; 1.

DR PROSITE; PS00002; SH3; 1.

DR SH2 domain; SH3 domain; Myristate; Phosphorylation;

DR Alternative splicing; Lipoprotein.

FT INIT MET 0

FT DOMAIN 21 81 SH3.

FT DOMAIN 83 174 SH2.

FT DOMAIN 189 280 SLA C-TERMINAL.

FT LIPID 1 N-myristoyl glycine.

Query Match 36.3%; Score 491.5; DB 1; Length 280;

Best Local Similarity 43.6%; Pred. No. 6.2e-34;

Matches 115; Conservative 38; Mismatches 82; Indels 29; Gaps 7;

Oy 9 KSLPSPS---LSSVQGGPVTWEAERSKATVALGSPFAGGPAELSLRLGELPTLVSED 65

Db 5 KSTSPSPERPLSSS-----EGLESDFLAV-LTDYSPDIPSPPIFRGEXLRVISDE 54

Oy 66 GDMWTVLSEVSGREYNI:PSVHVGVSHGVLVEGLSREKAEILLLPNPGGAFIRESOT 125

Db 55 GGMWKAISLTGRSEYIPGICVARVYHGMVFEGLGRDKAEILLQPDTKIGSPWIRESET 114

Oy 126 RRGYSYSVLRSPASPDNRHRYHCHLDNGWLYISPRITFPSQLVHDYSLADIDIC 185

Db 115 KKGFFVLSVR-----HRQVHKYRIFRLPNWYIISPRITFQCLDITVHYSEVADGLCC 168

Oy 196 LLKEPCVQLR-----AGPLGKGDIPLVTVORTPLNKKEDSSLLPSEAATG---EESL 236

Db 169 VLTPTCLAQNIPTATSPSPCTSGSPVTIRQKTFDKWRVSRLOEGSEGAEINPLRVDESIL 228

Oy 237 LSEGLRESLSFYIEL-NDEAVSLD 259

Db 229 PSYGLRESIASYLSLTGDDSSFD 252

RESULT 5

SLA1 HUMAN STANDARD; PR7; 275 AA.

ID SLA1\_HUMAN

AC Q13239; Q9UMQ8;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE SRC-like-adaptor (Src-like-adaptor protein 1) (hSLAP).

GN SLA OR SLAP OR SLAP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RN SEQUENCE FROM N.A., AND INDUCTION BY ATRA.

RP MEDLINE=96423054; PubMed=8825655;

RX Angriest M., Wells D.E., Chakravarti A., Pandey A.;

RT "Chromosomal localization of the mouse Src-like adaptor protein (Slap)

RT gene and its putative human homolog SLA.";

RL Genomics 30:623-625(1995).

RN [2]

RN SEQUENCE FROM N.A., AND INDUCTION BY ATRA.

RP TISSUE=Histocytic lymphoma;

RX MEDLINE=97148576; PubMed=9020066;

RA Ohtsuki T., Hatake K., Ikeda M., Tomizuka H., Terui Y., Uwai M.,

RA Miura Y.;

RT "Expression of Src-like adaptor protein mRNA is induced by all-trans

RT retinoic acid.";

RL Biochem. Biophys. Res. Commun. 230:81-84(1997).

RN [3]

RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RP TISSUE=Fetal brain;

RX MEDLINE=98321620; PubMed=9660183;

RA Meijerink P.H.S., Yanakiev P., Zorn I., Grierson A.J., Bikker H.,

RA Dye D., Kalaydjieva L., Baas F.;

RT "The gene for the human Src-like adaptor protein (hSLAP) is located

RT within the 64-kb intron of the thyroglobulin gene.";

RL Eur. J. Biochem. 254:297-303(1998).

RN [4]

RN SEQUENCE FROM N.A., AND FUNCTION.

RX MEDLINE=21553259; PubMed=11696592;

RA Holland S.J., Liac X.C., Mendenhall M.K., Zhou X., Pardo J., Chu P.,

RA Spencer C., Fu A.C., Sheng N., Yu P., Pali E., Nagin A., Shen M.,

RA Yu S., Chan E., Wu X., Li C., Wolsettschlag M., Aversa G.,

RA Kolbinger F., Bennett M.K., Molineaux S., Luo Y., Payan D.G.,

RA Mancebo H.S.Y., Wu J.;

RT "Functional cloning of Src-like adaptor protein-2 (SLAP-2), a novel

RT inhibitor of antigen receptor signaling.";

RL J. Exp. Med. 194:1263-1276(2001).

RN [5]

RN SEQUENCE FROM N.A.

RP TISSUE=Bone marrow;

RX MEDLINE=22389257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [6]

RN SEQUENCE OF 1-71 FROM N.A.

RP MEDLINE=21100465; PubMed=11179692;

RX Kratchmarova I., Sosinowski T., Weiss A., Witter K., Vincenz C.,

FT	MUTAGEN	217	217	ZAP70, CD3Z, SYK AND LAT.
FT				L->S: ABOLISHES INTERACTION WITH CBL.
FT				WHILE IT DOES NOT AFFECT DIMERIZATION
FT				WHEN ASSOCIATED WITH S-224 AND S-229.
FT	MUTAGEN	223		L->S: ABOLISHES INTERACTION WITH CBL.
FT				WHILE IT DOES NOT AFFECT DIMERIZATION
FT				WHEN ASSOCIATED WITH S-218 AND S-229.
FT	MUTAGEN	228		L->S: ABOLISHES INTERACTION WITH CBL.
FT				WHILE IT DOES NOT AFFECT DIMERIZATION
FT				WHEN ASSOCIATED WITH S-218 AND S-224.
FT	MUTAGEN	236		LSL->OSQ: ABOLISHES INTERACTION WITH CBL.
FT				SLIGHTLY AFFECTS DIMERIZATION.
FT	CONFLICT	70	70	Y -> D (IN REF. 6).
FT	SEQUENCE	275 AA;	31025 MW;	ED6837C99AD5534F CRC64;
Query Match		35.6%;	Score 481.5;	DB 1; Length 275;
Best Local Similarity		40.3%;	Pred. No. 4.1e-33;	
Matches 102;	Conservative	43;	Mismatches 85;	Indels 23; Gaps 4;
QY	9	KSLPSPSLSSVQGGPGVPTMEASRSKATAVALGSPAGGPAELSLRLGRLPLTVSDEGDW	68	
DB	5	KSTPAPA-----ERLPNPGLSDDFLAVLSDTPSPDISPPIFRGEGKLRVISDEGGW	57	
QY	69	WTVLSVSGREYNIPSVHGVKSHGWLVEGLSRKAEALLLPNGPGGAFLIRESQTRRG	128	
DB	58	WKAISLSTGRESYIPGICVARVTHGWLFGSLGRDKAEALLQLPDTKVGSGFMIRESETKKG	117	
QY	129	SVLSVLSRSPASWDRIHYRTHCLDNGWLYISPLTFPSLQALVDHYSELADDICLLK	198	
DB	118	FYSLSVR-----HRQVGHYRIFRPNPNWYYSPLTFQCLDELVNHYSEVADGLCCVLT	171	
QY	189	EPCLVQRAGPLPGKDIPLPVTVQTPPLNWKELDSLLFSEAATG-----EESLSSEGL	241	
DB	172	TFCLTQSTAPAVRASSPVLTRQKTVDRRYRVR---LQEDPGETENPLGVDSLSFSYGL	228	
QY	242	RESLSFYISLNDK 254		
DB	229	RESIASYLSLTSE 241		
RESULT 6				
LYN_HUMAN				
ID	LYN_HUMAN	STANDARD;	PRT;	511 AA.
AC	P07948;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Tyrosine-protein kinase LYN (EC 2.7.1.112).			
GN	LYN.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=87172710; PubMed=3561390;			
RX	Yamanashi Y., Fukushima S.-I., Samba K., Sukegawa J., Miyajima N.,			
RA	Matsubara K.-I., Yamamoto T., Toyoshima K.;			
RA	"The yes-related cellular gene lyn encodes a possible tyrosine kinase			
RT	similar to p56lck.";			
RT	Mol. Cell. Biol. 7:237-243(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=94171041; PubMed=8125304;			
RX	Rider L.G., Raben N., Miller L., Jelsema C.;			
RA	"The cDNAs encoding two forms of the LYN protein tyrosine kinase are			
RT	expressed in rat mast cells and human myeloid cells.";			
RT	Gene 138:219-222(1994).			
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein			
CC	tyrosine phosphate.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=LYN A;			



```

FT CONFLICT 431 L -> P (IN REF. 4).
SQ SEQUENCE 511 AA; 58681 MW; 3935222ICC90C50F0 CRC64;

Query Match 27.2%; Score 367.5; DB 1; Length 511;
Best Local Similarity 40.3%; Pred. NO. 3.2e-23;
Matches 81; Conservative 33; Mismatches 78; Indels 9; Gaps 3;

QY 6 SRKSLPSPSLSSVQCGPVTWMEASRSKATAVALGSPFAGGPAELSLRLGCEPLTIVSED 65
DB 37 SNKQPPVPPEF-HLLPQRFQTKDPEQGDIVVALYPDGIHPDDLFFKKGKMKVLEEH 95
QY 66 GDMWTVLSEVSGREYNIPSVHVQKV-----SHGWLVEGLSRKEAEILLLLDGNPGGAFLIR 121
DB 96 GEWMAKSLSSKREGIPSNYAKVNTLEETEEFFKDIKDAERQLLAPGNSAGAEFLIR 155
QY 122 ESOTREGSYSLSVLSRSPASMDRIHRVIRHCLDNGWLVIYISPRITFFPSLQALVDHYSELAD 181
DB 156 ESETLKGSFSLVRDYPMHGDVYKHKIRSLDNGGYISPRITFFPCISDMKHYKQSD 215
QY 182 DIICLLKEPCVLQAGPLPGK 202
DB 216 GLCRRLEKACI-----SPKPK 232

RESULT 8
LYN_RAT STANDARD; PRT; 511 AA.
ID LYN_RAT Q07014; Q63320; AC
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase LYN (EC 2.7.1.112).
OS LYN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
[1]
SEQUENCE FROM N.A.
RA Minoguchi K., Nishikata H., Siraganian R.P.;
RT "Bacterially expressed rat p56lyn binds several proteins in rat
RT basophilic leukemia cells including pp72, a tyrosine phosphorylated
RT protein prominent in activated cells";
RL J. Immunol. 150:222-222(1993).
[2]
SEQUENCE FROM N.A.
RA MEDLINE=94171041; PubMed=8125304;
RA Rider L.G., Raben N., Miller L., Jelsema C.;
RT "The cDNAs encoding two forms of the LYN protein tyrosine kinase are
RT expressed in rat mast cells and human myeloid cells.";
RL Gene 138:219-222(1994).
[3]
SEQUENCE FROM N.A.
RA MEDLINE=97442484; PubMed=9295361;
RA Vonakis B.M., Chen H., Haleem-Smith H., Metzger H.;
RT "The unique domain as the site on lyn kinase for its constitutive
RT association with the high affinity receptor for Ige.";
RL J. Biol. Chem. 272:24072-24080(1997).
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Names=LYN A;
CC IsoIdc=Q07014-1; Sequence=Displayed;
CC Names=LYN B;
CC IsoIdc=Q07014-2; Sequence=VSP_005004;
CC -I- TISSUE SPECIFICITY: Expressed predominantly in B-lymphoid and
CC myeloid cells.
CC -I- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC
CC subfamily.
CC -I- SIMILARITY: Contains 1 SH2 domain.
CC -I- SIMILARITY: Contains 1 SH3 domain.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L14951; AAA41549.1; -  
EMBL; L14782; AAA20944.1; -  
EMBL; L14823; AAA20945.1; -  
EMBL; AF000300; AAB71344.1; -  
EMBL; AF000301; AAB71345.1; -  
EMBL; AF000302; AAB71346.1; -  
PIR; I56360; I56160.  
HSSP; P08631; IAD5.  
InterPro; IPR000719; Prot\_kinase.  
InterPro; IPR000980; SH2.  
InterPro; IPR001452; SH3.  
InterPro; IPR001245; Tyr\_kinase.  
InterPro; IPR008266; Tyr\_kinase\_AS.  
Pfam; PF00069; Pkinase; 1.  
Pfam; PF00017; SH2; 1.  
Pfam; PF00018; SH3; 1.  
PRINTS; PR00401; SH2DOMAIN.  
PRINTS; PR00452; SH3DOMAIN.  
PRINTS; PR00109; TYRKINASE.  
ProDom; PD000001; Prot\_kinase; 1.  
ProDom; PD000093; SH2; 1.  
ProDom; PD000066; SH3; 1.  
SMART; SM00252; SH2; 1.  
SMART; SM00326; SH3; 1.  
SMART; SM00219; TyrKc; 1.  
PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
PROSITE; PS50001; SH2; 1.  
PROSITE; PS50002; SH3; 1.  
Proto-oncogene; Tyrosine-protein kinase; Phosphorylation;  
Transferase; ATP-binding; Myristate; SH2 domain; SH3 domain;  
Palmitate; Lipoprotein; Alternative splicing.  
INIT MET 0  
LIPID 1 1 N-myristoyl glycine (By similarity).  
LIPID 2 2 S-palmitoyl cysteine (By similarity).  
DOMAIN 62 122 SH3.  
DOMAIN 128 225 SH2.  
DOMAIN 246 500 PROTEIN KINASE.  
NP\_BIND 252 260 ATP (BY SIMILARITY).  
BINDING 274 274 ATP (BY SIMILARITY).  
ACT\_SITE 366 366 BY SIMILARITY.  
MOD\_RES 396 396 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
MOD\_RES 507 507 PHOSPHORYLATION (BY SIMILARITY).  
VARSPIC 24 44 Missing (in isoform LYN B).  
CONFLICT 230 230 /FTID=VSP\_005004.  
CONFLICT 307 307 P -> L (IN REF. 2).  
CONFLICT 418 418 V -> A (IN REF. 2).  
SEQUENCE 511 AA; 58529 MW; 24425E5E229CD43ED CRC64;  
Query Match 27.2%; Score 367.5; DB 1; Length 511;  
Best Local Similarity 40.3%; Pred. No. 3.2e-23;  
Matches 81; Conservative 33; Mismatches 78; Indels 9; Gaps 3;  
QY 6 SRRKSLPSLSSVQGGPVTMEATSKATAVALGSPGPAELSLRLGELPTIVSED 65  
DB 37 SNKQRPYPE-SQLPGQRFQAKDFEQGDIVVALPYDGIHPDLSFKKGEKMKVLEH 95  
QY 66 GDMWTVLSVSGREYNIPSVHGVK-----SHGWLVEGLSRKAELLILLPNPGAFILR 121  
DB 96 GSWAKAKSLSSREGFIPSNVYKNTLETETWEFFKIDTRKDAERQLLAPGNSAGAFILR 155  
QY 122 ESETRGYSLSVLSRSPASWDRIRHYRTHCLDNGWLYISPLTPPSLQALVDHYSELAD 181

Db 156 ESETLKGSFSLSVRDYPMHGDVILKHYKIRSLDNGGYISPRITFCISDMIKHYQKSD 215  
QY 182 DICLLKEPCVLQAGPLPK 202  
Db 216 GLCRRLEKACI-----SPKPK 232  
RESULT 9  
HCK\_HUMAN  
ID HCK\_HUMAN STANDARD; PRT; 525 AA.  
AC P08631; Q96CC0; Q9HSY5; Q9NUA4; Q9UMJ5;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tyrosine-protein kinase HCK (BC 2.7.1.112) (p59-HCK/p60-HCK)  
DE (Hemopoietic cell kinase).  
GN HCK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE OF 21-525 FROM N.A.  
RX MEDLINE=87257942; PubMed=3496523;  
RA Quintrell N., Lebo R., Varmus H., Bishop J.M., Pettenati M.J.,  
le Beau M.M., Diaz M.O., Rowley J.D.; that encodes a protein-tyrosine  
RT "Identification of a human gene (HCK) that encodes a protein-tyrosine  
RT kinase and is expressed in hemopoietic cells."  
RL Mol. Cell. Biol. 7:2267-2275(1987).  
RN [2]  
RP SEQUENCE OF 21-525 FROM N.A.  
RX MEDLINE=87257943; PubMed=3453117;  
RA Ziegler S.F., March J.D., Lewis D.B., Perlmuter R.M.;  
RT "Novel protein-tyrosine kinase gene (hck) preferentially expressed in  
RT cells of hematopoietic origin."  
RL Mol. Cell. Biol. 7:2276-2285(1987).  
RN [3]  
RP SEQUENCE OF 21-525 FROM N.A.  
RC TISSUE-B-cell.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K.,  
Holtkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Rappaport J.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Udwin T.B., Toshyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
Bosak S.A., McEwan P.J., McKernan K.G., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 21-525 FROM N.A.  
RC TISSUE=ileal mucosa;  
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,  
Yamada K., Fujii Y., Ozaki K., Hiraiz M., Ohmori Y., Ota T., Suzuki Y.,  
Oyashiki M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,  
Isoigai T., Sugano S.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.V., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor P.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dunn P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.D., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Leinvaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McInnes J.C., Nickerson T.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
RA Stuce C.D., Smith M.L., Soderlund C., Stewart C.A., Suleston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,  
RA Whitehead S.L., Whittaker P., Willey D.I., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
Rogers J.,  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RT Nature 414:865-871(2001).  
RN [6]  
RP SEQUENCE OF 178-525 FROM N.A.  
RX TISSUE=Splicein;  
RC MEDLINE=92241680; PubMed=1572549;  
RA Hradetzky D., Streibhardt K., Rueaemen-Waigmann H.;  
RT "The genomic locus of the human hemopoietic-specific cell protein  
RT tyrosine kinase (PTK)-encoding gene (HCK) confirms conservation of  
RT exon-intron structure among human PTKs of the src family.";  
RL Gene 113:275-280(1992).  
RN [7]  
RP SEQUENCE OF 1-21 FROM N.A., AND ALTERNATIVE INITIATION.  
RX MEDLINE=91342636; PubMed=1875927;  
RA Lock P., Ralph S., Stanley E., Boulet I., Ramsay R., Dunn A.R.;  
RT "Two isoforms of murine hck, generated by utilization of alternative  
RT translational initiation codons, exhibit different patterns of  
RT subcellular localization.";  
RL Mol. Cell. Biol. 11:4363-4370(1991).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 77-525.  
RX MEDLINE=97171106; PubMed=9024658;  
RA Sighieri F., Moaresi I., Kuriyan J.;  
RT "Crystal structure of the Src family tyrosine kinase Hck.";  
RL Nature 385:602-609(1997).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 80-136.  
RX MEDLINE=98453315; PubMed=9778343;  
RA Arold S., O'Brien R., Franken P., Strub M.P., Hoh F., Dumas C.,  
RA Ladbury J.E.;  
RT "A loop flexibility enhances the specificity of Src family SH3  
RT domains for HIV-1 Nef.";  
RL Biochemistry 37:14683-14691(1998).  
RN [10]  
RP STRUCTURE BY NMR OF 77-137.  
RX MEDLINE=98239731; PubMed=9571048;  
RA Horita D.A., Baldisseri D.M., Zhang W., Altieri A.S., Smithgall T.E.,  
RA Gmeiner W.H., Byrd R.A.;  
RT "Solution structure of the human Hck SH3 domain and identification of  
RT its ligand binding site.";  
RL J. Mol. Biol. 278:253-285(1998).  
RN [11]  
RP STRUCTURE BY NMR OF 138-244.  
RX MEDLINE=97263487; PubMed=9109402;  
RA Zhang W., Smithgall T.E., Gmeiner W.H.;  
RT "Sequential assignment and secondary structure determination for the  
RT Src homology 2 domain of hematopoietic cellular kinase.";  
RL FEBS Lett. 406:131-135(1997).

CC -!- FUNCTION: May serve as part of a signaling pathway coupling the Fc  
CC receptor to the activation of the respiratory burst. May also  
CC contribute to neutrophil migration and may regulate the  
CC degranulation process of neutrophils.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Isoform p60-HCK and isoform p59-HCK are  
CC associated with membranes. Isoform p60-HCK is also cytoplasmic (By  
CC similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative initiation;  
CC Comment=2 isoforms, p60-HCK (shown here) and p59-HCK, are  
CC produced by alternative initiation;  
CC -!- TISSUE SPECIFICITY: Expressed predominantly in cells of the  
CC myeloid and b-lymphoid lineages  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M16591; AAA52643.1; -;  
CC EMBL: M16592; AAA52644.1; -;  
CC EMBL: BC014435; AAH14435.1; -;  
CC EMBL: AK026432; BAB15482.1; -;  
CC EMBL: AL049539; CAB75606.1; -;  
CC EMBL: X58741; CAA41565.2; -;  
CC EMBL: X58742; CAA41565.2; JOINED.  
CC EMBL: X58743; CAA41565.2; JOINED.  
CC PIR: A27811; TVHUHC  
CC PDB: 2HCK; 20-AUG-97.  
CC PDB: 3HCK; 15-OCT-97.  
CC PDB: 4HCK; 17-JUN-98.  
CC PDB: 5HCK; 17-JUN-98.  
CC PDB: 1AD5; 15-MAY-97.  
CC PDB: 1BUI; 11-NOV-98.  
CC PDB: 1OCF; 08-JUN-99.  
CC Genew; HGNC:4840; HCK.  
CC MIM; 142370; -;  
CC GO: GO:0004713; F:protein-tyrosine kinase activity; TAS.  
CC GO: GO:0007498; P:mesoderm development; TAS.  
CC GO: GO:0006468; P:protein amino acid phosphorylation; TAS.  
CC InterPro: IPR000719; Prot\_kinase.  
CC InterPro: IPR000980; SH2.  
CC InterPro: IPR001452; SH3.  
CC InterPro: IPR001245; Tyr\_kinase.  
CC InterPro: IPR008266; Tyr\_kinase\_AS.  
CC Pfam: PF00069; kinase; 1.  
CC Pfam: PF00017; SH2; 1.  
CC Pfam: PF00018; SH3; 1.  
CC PRINTS; PR00401; SH2DOMAIN.  
CC PRINTS; PR00452; SH3DOMAIN.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom: PD000001; Prot\_kinase; 1.  
CC ProDom: PD000093; SH2; 1.  
CC ProDom: PD000066; SH3; 1.  
CC ProDom: PD000066; SH3; 1.

Query Match 26.6%; Score 360.5; DB 1; Length 525;  
Best Local Similarity 41.6%; Pred. No. 1.3e-22;  
Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;

QY 12 PPSLSVVGQGGPVTMEAKRSKATAVALGSPAGGPAELSLRLGELPTIVSDGQWTV 71

Db 60 PGNSHNS---NTPGIREAGSEDIIVVALDYEAHHDLSFKGDQWVYVLESGSWKA 116

QY 72 LSEVSGREYNIPSVHVGVK-----SHGWLYEGLSRKAEELLLPGNPGGAFLIRSQTR 127

DR PROSITE: P550002; SH3: 1.  
KW Transferase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;  
KW Myristate; SH2 domain; SH3 domain; Lipoprotein.  
FT INIT MET 0 0  
FT DOMAIN 55 115 SH3.  
FT DOMAIN 121 218 SH2.  
FT DOMAIN 239 492 PROTEIN KINASE.  
FT NP BIND 245 253 ATP (BY SIMILARITY).  
FT BINDING 267 267 ATP (BY SIMILARITY).  
FT ACT SITE 358 358 ATP (BY SIMILARITY).  
FT LIPID 1 1 N-myristoyl glycine (By similarity).  
FT MOD\_RES 388 388 PHOSPHORYLATION (AUTO-), (BY SIMILARITY).  
FT SEQUENCE 503 AA; 56964 MW; 561F9322D2DE3436 CRC64;  
Query Match 26.5%; Score 358; DB 1; Length 503;  
Best Local Similarity 45.3%; Pred. NO. 2e-22;  
Matches 72; Conservative 29; Mismatches 54; Indels 4; Gaps 1;  
QY 38 VALGSPAGGPAELSLRLGELPTIVSEGDGWTVLSEVSGREYNIPSVHGVK----SHG 93  
DB 61 VALDYTEAITHHDSLPQKGDQWVLEESGWNKARSLATKKGYSINVARVDSLETEE 120  
QY 94 WLYEGHSREKAEPELLLPNGPGAFILRSQTRRGYSLSVLRSPASWDRIHYRIHCL 153  
DB 121 WFFKGISRKDAERQLLAPGNMLGSFMRDSETTKGSYLSVRDYPDPRQGTVKHYKIRL 180  
QY 154 DNGWLVISPLRFPSQALVDHYSELADDCILLKEPCV 192  
DB 181 DNGGFIISPRSFSTLQELVDHYKSGDGLCQKLSVPCV 219  
RESULT 11  
BLK\_HUMAN STANDARD; PRT; 504 AA.  
ID BLK\_HUMAN AC P51451; Q16291;  
AC 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DT Tyrosine-protein kinase BLK (EC 2.7.1.112) (B lymphocyte kinase) (p55-  
DE BLK).  
DE GN.  
OS Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RP MEDLINE=95123078; PubMed=7822795;  
RA Drebin J.A., Hartzell S.W., Griffin C., Campbell M.J.,  
RA Niederhuber J.E.;  
RA "Molecular cloning and chromosomal localization of the human homologue  
RT of a B-lymphocyte specific protein tyrosine kinase (blk).";  
RL Oncogene 10:477-485(1995).  
CC -!- FUNCTION: May function in a signal transduction pathway that is  
CC restricted to B lymphoid cells.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z33998; CA883965.1; -
CC EMBL; S76617; AAB3265.1; -
CC PIR; I37206; I37206.
CC HSP; P16277; 1BLK.
CC Genew; HGNC:1057; BLK.
CC MIM; 191305; -
CC GO; GO:0004713; F:protein-tyrosine kinase activity; TAS.
CC GO; GO:0007243; P:protein kinase cascade; TAS.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR008266; Tyr_kinase_AS.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00017; SH2; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00401; SH2DOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC ProDom; PD000093; SH2; 1.
CC ProDom; PD000066; SH3; 1.
CC SMART; SM00252; SH2; 1.
CC SMART; SM00326; SH3; 1.
CC SMART; SM0219; Tyrc; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00109; PROTEIN KINASE_TYR; FALSE_NEG.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00001; SH2; 1.
CC PROSITE; PS00002; SH3; 1.
CC Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding;
KW Myristate; SH2 domain; SH3 domain; Lipoprotein.
FT INIT_MET 0 0
FT LIPID 1 1
FT DOMAIN 57 117
FT DOMAIN 123 219
FT DOMAIN 240 493
FT NP_BIND 246 254
FT BINDING 268 268
FT ACT_SITE 359 359
FT MOD_RES 388 388
FT CONFLICT 286 286
FT CONFLICT 406 406
SQ SEQUENCE 504 AA; 57607 MW; BD1DF50EC7370C8 CRC64;
Query Match 26.1%; Score 352.5; DB 1; Length 504;
Best Local Similarity 43.6%; Pred. No. 5,7e-22;
Matches 75; Conservative 24; Mismatches 58; Indels 5; Gaps 2;
QY 25 PVTWEAKSKATAVALGSPFAGGPAELSLRIGLEPLTIVSDGDWVTVLSVSGREYNIPS 84
Db 50 PPDEHLDKHFVVALDYDTAMNDRLQMLKGEKQLVKGTDGWLARSILVTGREGVPS 109
QY 85 VHVGVKVS---HGWLVEGLSKAEAEILLPGNPGGAFLIRESGTSGYSLSVLSRPA 140
Db 110 NFVARVESLEMERWFFRQGRKEARQLLAPINKAGSFLIRESETNKGAFSLVK-DVTT 168
QY 141 SWDIRHYRIHCLDNGWLYISPRITFTPSLQALVDHYSELADDICLLKEPCV 192
Db 169 QGELIKHYKIRCLDGGYIISPRITFTPSLQALVDHYSELADDICLLKEPCV 220
RESULT 12
HCK_MOUSE
ID HCK_MOUSE
AC P08103
DT 01-AUG-1988 (Rel. 08, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
```

```
DE Tyrosine-protein kinase HCK (BC 2.7.1.112) (p56-HCK/p59-HCK)
DE (Hemopoietic cell kinase) (B-cell/myeloid kinase) (BMK).
GN HCK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE OF 21-523 FROM N.A.
RC STRAIN=ICR; TISSUE=Macrophage;
RX MEDLINE=88067781; PubMed=3684607;
RA Klemz M.J., McRercher S.R., Maki R.A.;
RT "Nucleotide sequence of the mouse hck gene.";
RL Nucleic Acids Res. 15:9600-9600(1987).
RN [2]
RN SEQUENCE OF 21-523 FROM N.A.
RX MEDLINE=88068587; PubMed=3317404;
RA Holtzman D.A., Cook W.D., Dunn A.R.;
RT "Isolation and sequence of a cDNA corresponding to a src-related gene
RT expressed in murine hemopoietic cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8325-8329(1987).
RN [3]
RN SEQUENCE OF 1-21 FROM N.A., AND ALTERNATIVE INITIATION.
RX MEDLINE=91342636; PubMed=1875927;
RA Lock P., Ralph S., Stanley E., Boulet J., Ramsay R., Dunn A.R.;
RT "Two isoforms of murine hck, generated by utilization of alternative
RT translational initiation codons, exhibit different patterns of
RT subcellular localization.";
RL Mol. Cell. Biol. 11:4363-4370(1991).
CC -!- FUNCTION: May serve as part of a signaling pathway coupling the Fc
CC receptor to the activation of the respiratory burst. May also
CC contribute to neutrophil migration and may regulate the
CC degranulation process of neutrophils.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Isoform p59-HCK and isoform p56-HCK are
CC associated with membranes. Isoform p59-HCK is also cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, p59-HCK (shown here) and p56-HCK, are
CC produced by alternative initiation;
CC -!- TISSUE SPECIFICITY: Expressed predominantly in cells of the
CC myeloid and b-lymphoid lineages.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC
CC subfamily.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y00487; CAA68544.1; -
CC EMBL; J03023; AAA37305.1; -
CC PIR; A27282; TWGSHC.
CC HSP; P08631; 1AD5.
CC MGD; MG1:96052; HCK.
CC InterPro; IPR00108; New_cyt_fact_2.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR008266; Tyr_kinase_AS.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00017; SH2; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00499; P67PHOX.
CC PRINTS; PR00401; SH2DOMAIN.
CC PRINTS; PR00452; SH3DOMAIN.
```

Okano Y., Sugimoto Y., Fukuoka M., Matsui A., Nagata K.I., Nozawa Y.;  
"Identification of rat cDNA encoding hck tyrosine kinase from  
megakaryocytes."; *Biochem. Biophys. Res. Commun.* 181:1137-1144 (1991).  
[2]  
SEQUENCE FROM N.A.  
SPRAIN=Wistar; TIGSUS=Spleen;  
Vijaya Gouri B.S., Renu V., Kamatkar S., Swarup G.;  
"Nucleotide sequence of a cDNA coding for rat hck tyrosine kinase and  
characterization of its gene product."; *J. Biosci.* 19:117-129 (1994).  
CC -1- FUNCTION: May serve as part of a signaling pathway coupling the Fc  
receptor to the activation of the respiratory burst. May also  
contribute to neutrophil migration and may regulate the  
degranulation process of neutrophils.  
CC CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated.  
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
subfamily.  
CC -1- SIMILARITY: Contains 1 SH2 domain.  
CC -1- SIMILARITY: Contains 1 SH3 domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
EMBL; S74141; AAB20754.1; --  
EMBL; M83666; AAA41312.1; --  
EMBL; X62345; CAA44218.1; --  
PIR; JQ1321; JQ1321.  
HSP; P08631; IEU1.  
InterPro; IPR000719; Prot\_kinase.  
InterPro; IPR000980; SH2.  
InterPro; IPR001452; SH3.  
InterPro; IPR001245; Tyr\_kinase.  
InterPro; IPR008266; Tyr\_kinase\_AS.  
Pfam; PF00069; pkinase; 1.  
Pfam; PF00017; SH2; 1.  
Pfam; PF00018; SH3; 1.  
PRINTS; PR00401; SH2DOMAIN.  
PRINTS; PR00452; SH3DOMAIN.  
PRINTS; PR00109; TYRKINASE.  
ProDom; PD000001; Prot\_kinase; 1.  
ProDom; PD000093; SH2; 1.  
ProDom; PD000066; SH3; 1.  
SMART; SMO0252; SH2; 1.  
SMART; SMO0326; SH3; 1.  
SMART; SMO0219; Tyrcg; 1.  
PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
PROSITE; PS50001; SH2; 1.  
PROSITE; PS50002; SH3; 1.  
Transferase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;  
Myristate; SH2 domain; SH3 domain; Lipoprotein.  
INIT NET 0  
DOMAIN 54 114  
FT DOMAIN 120 217  
FT DOMAIN 238 491  
FT NP BIND 244 252  
FT BINDING 266 266  
FT ACT SITE 357 357  
FT LIPID 1 1  
FT MOD RES 387 387  
FT CONFLICT 50 50  
FT CONFLICT 204 204  
FT CONFLICT 305 305  
FT CONFLICT 502 AA; 56885 MW; 4CFC1F3F0E82EADF CRC64;  
SQ SEQUENCE

Query Match 26.0%; Score 351.5; DB 1; Length 502;  
Best Local Similarity 40.9%; Pred. No. 6.9e-22;  
Matches 83; Conservative 32; Mismatches 73; Indels 15; Gaps 4;

QY 5 PSRRKSLPSLSSVQCGPVTMEARSKAT-AVALGSPAGPAELSLRLGEPITIVS 63  
DB 32 PTPSPKLGPSINSLPG-----FVSGSEDTIVALDYEAHREDLSFKQGDQMVLE 85  
QY 64 EDGQWTVLSEVSGREYNIPSVHVGVKVS-----SHGWLVEGLSRKAELELLLPNGPGAPL 119  
DB 86 ESGWVKARSLSATKEGVISYVAVNSIETEWFPGISRKDAERHLLAPGNMLGSEFM 145  
QY 120 IRESQTRGSGYSLVRLSPASMDRIRHYRHICLDNGWLYISPRLTTPSPSQALVDHYSEL 179  
DB 146 IRDSETHKGSVLSVRDFDPQHGDTVXHYKIRTLDSGGFVISPSTFSSIQELVWYKKG 205  
QY 180 ADDICLLKEPCVLRAGPLPGK 202  
DB 206 KDLGCKLSVPCV-----SPKPK 224

RESULT 14  
BLK\_MOUSE STANDARD; PRT; 498 AA.

AC P16277;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tyrosine-protein kinase BLK (BC 2.7.1.112) (B lymphocyte kinase) (p55-  
BLK).  
GN BLK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=B-cell;  
RX MEDLINE=90117147; PubMed=2404338;  
RA Dymekki S.M., Niederhuber J.E., Desiderio S.V.;  
RT "Specific expression of a tyrosine kinase gene, blk, in B lymphoid  
cells.";  
RL Science 247:332-336(1990).  
RN [2]  
RP STRUCTURE BY NMR OF SH2 DOMAIN.  
RX MEDLINE=96224819; PubMed=8639560;  
RA Metzler W.J., Leitinger B., Pryor K., Mueller L., Farmer B.T. II;  
RT "The three-dimensional solution structure of the SH2 domain from  
p55blk kinase.";  
RL Biochemistry 35:6201-6211(1996).  
CC -!- FUNCTION: May function in a signal transduction pathway that is  
restricted to B lymphoid cells.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
tyrosine phosphate.  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. SRC  
subfamily.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
-----  
EMBL; M30903; AAA40453.1; -.  
DR PIR; A40092; A40092.  
DR PDB; 1BLJ; 12-MAR-97.  
DR PDB; 1BLK; 12-MAR-97.  
DR MGD; MGI:88169; BLK.

DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001452; Tyr\_kinase.  
DR InterPro; IPR001452; Tyr\_kinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR ProDom; PD000093; SH2; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00219; TyKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; FALSE\_NEG.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00001; SH2; 1.  
DR PROSITE; PS00002; SH3; 1.  
KW Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding;  
KW Myristate; SH2 domain; SH3 domain; 3D-structure; Lipoprotein.  
FT INIT\_MET 0 0  
FT LIPID 1 1 N-myristoyl glycine (By similarity).  
FT DOMAIN 51 111 SH2.  
FT DOMAIN 117 213 SH3.  
FT NP\_BIND 234 487 PROTEIN KINASE.  
FT BINDING 262 262 ATP (BY SIMILARITY).  
FT ACT\_SITE 353 353 ATP (BY SIMILARITY).  
FT MOD\_RES 382 382 BY SIMILARITY.  
FT HELIX 124 131 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT TURN 132 132  
FT TURN 134 135  
FT TURN 138 139  
FT STRAND 141 145  
FT TURN 147 148  
FT STRAND 153 157  
FT STRAND 159 161  
FT TURN 162 163  
FT STRAND 164 166  
FT STRAND 170 175  
FT TURN 176 178  
FT STRAND 179 183  
FT TURN 184 185  
FT STRAND 186 188  
FT HELIX 191 197  
FT TURN 198 200  
SQ SEQUENCE 498 AA; 56513 MW; BB49D7B079FDD577 CRC64;

Query Match 25.2%; Score 341.5; DB 1; Length 498;  
Best Local Similarity 40.0%; Pred. No. 4.7e-21;  
Matches 74; Conservative 29; Mismatches 65; Indels 17; Gaps 3;

QY 12 PPSLSVSSVQCGPVTMEARSKATAVAGSPAGPAELSLRLGEPITIVSGDWTV 71  
DB 43 PPSN-----QDPDEERFVVALDYAAVNDRLQVLKGEKQLVLRSTGDMWLA 90  
QY 72 LSEVSGREYNIPSVHVGVKVS-----HGWLYEGLSRKAELELLLPNGPGAPLIRSQTR 127  
DB 91 RSLVTRGEGVPSNFVAPVETLEVKWFFRTSRKDAERQLLAPMKAGSFLIRSESNK 150  
QY 128 GSYLSVRLSPASMDRIRHYRHICLDNGWLYISPRLTTPSPSQALVDHYSELADICLL 187  
DB 151 GAPSLSVK-DITTOGQGVVVKHYKIRSLDNGGYISPRITPTTLQALVQHYSKKGDLGCKL 209  
QY 188 KEPCV 192  
DB 210 TLPCV 214



Search completed: February 20, 2004, 09:30:30  
Job time : 41 secs

RA Bergman M., Mustelin T., Oetken C., Partanen J., Flint N.A.,  
RA Amrein K.E., Autero M., Burn P., Altalo K.;  
RT "The human p50csk tyrosine kinase phosphorylates p56lck at Tyr-505 and  
RT down regulates its catalytic activity.";  
RL EMBO J. 11:2919-2924(1992).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 52-225.  
RX MEDLINE=94203291; PubMed=7512222;  
RA Eck W.J., Atweell S.K., Shoelson S.E., Harrison S.C.;  
RT "Structure of the regulatory domains of the Src-family tyrosine  
RT kinase Lck.";  
RL Nature 368:764-769(1994).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 126-220.  
RX MEDLINE=95173978; PubMed=7332720;  
RA Mikol V., Baumann G., Keller T.H., Manning U.M., Zurini M.G.M.;  
RT "The crystal structures of the SH2 domain of p56lck complexed with  
RT two phosphopeptides suggest a gated peptide binding site.";  
RL J. Mol. Biol. 246:344-355(1995).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (1.0 ANGSTROMS) OF 121-225.  
RX MEDLINE=96177765; PubMed=8604142;  
RA Tong L., Warren T.C., King J., Betageri R., Rose J., Jakes S.;  
RT "Crystal structures of the human p56lck SH2 domain in complex with  
RT two short phosphotyrosyl peptides at 1.0-A and 1.8-A resolution.";  
RL J. Mol. Biol. 256:601-610(1996).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 118-225.  
RX MEDLINE=98352059; PubMed=9685372;  
RA Tong L., Warren T.C., Lukas S., Schembri-King J., Betageri R.,  
RA Proudfoot J.R., Jakes S.;  
RT "Carboxymethyl-phenylalanine as a replacement for phosphotyrosine in  
RT SH2 domain binding.";  
RL J. Biol. Chem. 273:20238-20242(1998).  
RN [20]  
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 230-500.  
RX MEDLINE=97100952; PubMed=8945479;  
RA Yamaguchi H., Hendrickson W.A.;  
RT "Structural basis for activation of human lymphocyte kinase Lck upon  
RT tyrosine phosphorylation.";  
RL Nature 384:484-489(1996).  
CC -!- FUNCTION: May participate in antigen-induced T-cell activation.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBUNIT: Binds to phosphatidylinositol 3'-kinase (PI3K) from T  
CC lymphocytes through its SH3 domain and to the tyrosine  
CC phosphorylated form of KDRBS1/p70 through its SH2 domain.  
CC -!- SUBCELLULAR LOCATION: BOUND TO THE CYTOPLASMIC DOMAIN OF EITHER  
CC CD4 OR CD8.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long;  
CC IsoId=P06239-1; Sequences=Displayed;  
CC Name=Short;  
CC IsoId=P06239-2; Sequences=VSP 005000, VSP 005001;  
CC Note=No experimental confirmation available;

Query Match 25.1%; Score 340; DB 1; Length 508;  
Best Local Similarity 40.6%; Pred. No. 6.4e-21;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
QY 26 VTWEAERSKAT-----AVALGSPAGGPAELSLRLGELPLTIVSDGDWTVLSEVSGRE 79  
DB 48 VTVEGSPASPQLQDNLVIALHSYEPHDGLDPEKGEQLRIEQSGEWKQAQSLTTGQE 107  
QY 80 YNIPSVHVGKVS-----HGWLYEGLGREKAEELLILPGNPGGAFLIREQTRGYSLSVR 135  
DB 108 GFIPNFVAKANSLPEPEPFFKNLSRKABRQLAPGNTHGSLFIRESESTAGSFSLSVR 167  
QY 136 LSRPASWDRIYRIVHCLDNGWLYISPLTPPSLQALVDHYSELADDICCLIKPCVQLQR 195  
DB 168 DFDQNGEVVKYKYNLDNGGFIYISPRITFFGLHVLVRYHYNASDGLCTLRSRCQTK 227

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:27:57 ; Search time 45 Seconds  
(without alignments)  
299.431 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSILPSRKSLSPSLSSV.....RESLFSYISLNDRAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	370.5	27.4	512	4	US-08-426-509A-16
2	370.5	27.4	512	4	US-08-232-545-16
3	370.5	27.4	512	5	PCT-US95-05008-16
4	360.5	26.6	505	4	US-08-426-509A-17
5	360.5	26.6	505	4	US-08-232-545-17
6	360.5	26.6	505	5	PCT-US95-05008-17
7	344.5	25.5	499	4	US-08-426-509A-19
8	344.5	25.5	499	4	US-08-232-545-19
9	344.5	25.5	499	5	PCT-US95-05008-19
10	340	25.1	508	4	US-09-862-154-1
11	340	25.1	509	3	US-09-033-555B-17
12	340	25.1	509	4	US-08-426-509A-18
13	340	25.1	509	4	US-09-457-040B-8
14	340	25.1	509	4	US-08-232-545-18
15	340	25.1	509	5	PCT-US95-05008-18
16	315.5	23.3	537	4	US-08-426-509A-11
17	315.5	23.3	537	4	US-08-232-545-11
18	315.5	23.3	537	5	PCT-US95-05008-11
19	315.5	23.3	543	4	US-08-426-509A-14
20	315.5	23.3	543	4	US-08-232-545-14
21	315.5	23.3	543	4	US-09-470-881-8
22	315.5	23.3	543	5	PCT-US95-05008-14
23	313.5	23.2	496	2	US-09-006-675-2
24	313.5	23.2	496	3	US-09-228-603A-2
25	312.5	23.1	529	4	US-08-426-509A-15
26	312.5	23.1	529	4	US-08-232-545-15
27	312.5	23.1	529	5	PCT-US95-05008-15

28	301	22.2	536	4	US-08-426-509A-12	Sequence 12, Appl
29	301	22.2	536	4	US-08-232-545-12	Sequence 12, Appl
30	301	22.2	536	5	PCT-US95-05008-12	Sequence 12, Appl
31	286.5	21.2	533	1	US-07-820-011A-2	Sequence 2, Appl
32	286.5	21.2	533	4	US-09-470-881-3	Sequence 3, Appl
33	286.5	21.2	533	5	PCT-US93-00445-2	Sequence 2, Appl
34	283.5	21.0	532	1	US-08-594-447-1	Sequence 1, Appl
35	283.5	21.0	532	2	US-08-665-647-1	Sequence 1, Appl
36	278	20.5	98	2	US-08-473-078-7	Sequence 7, Appl
37	276.5	20.4	536	1	US-07-820-011A-4	Sequence 4, Appl
38	276.5	20.4	536	4	US-08-426-509A-13	Sequence 13, Appl
39	276.5	20.4	536	4	US-08-232-545-13	Sequence 13, Appl
40	276.5	20.4	536	5	PCT-US93-00445-4	Sequence 4, Appl
41	276.5	20.4	536	5	PCT-US95-05008-13	Sequence 13, Appl
42	272	20.1	98	2	US-08-479-078-6	Sequence 6, Appl
43	266	19.7	108	5	PCT-US94-01840-6	Sequence 6, Appl
44	264	19.5	101	2	US-08-574-959A-5	Sequence 5, Appl
45	264	19.5	101	3	US-09-357-014-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-08-426-509A-16  
; Sequence 16, Application US/08426509A  
; Patent No. 6326469  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Gishizky, Mikhail  
; APPLICANT: Sures, Irman G.  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
; TITLE OF INVENTION: TYROSINE KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York,  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/426,509A  
; FILING DATE: 21-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/232,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-0074-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 512 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: No. 6326469e  
US-08-426-509A-16

Query Match 27.4%; Score 370.5; DB 4; Length 512;  
Best Local Similarity 39.8%; Pred. No. 1.7e-31;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY 6 SRRKSLPSLSVSSVQGGPVMTMEARSKATAVALGSPAGGPAELSLRLGEPLTIVSED 65  
 DB 38 SNKQRPVE-SQLPGQRFQTKDPEEQGDIVVALYPYDGIHPDDLSPFKGKMKVLEEH 96  
 QY 66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKABELLLPKNPGGAFLIR 121  
 DB 97 GEWMKAKSLTKKEGFIPSNYAKLNTLETEWFFKDIKDAERQQLAPGNSAGAFILR 156  
 QY 122 ESQTRGSSVLSVRLSRPASWDRIHRHICLDNGWLYISPRLTFFSLQALVDHYSELAD 181  
 DB 157 ESETLKGSSVLSVRDPDPVHGDIKIRSLDNGGYIISPRITPCISDMIKHYKQAD 216  
 QY 182 DICCLKEPCVLRAGPLPK 202  
 DB 217 GLCRLKACI-----SPKPK 233  
 RESULT 2  
 US-08-232-545-16  
 ; Sequence 16, Application US/08232545  
 ; Patent No. 6506578  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ulrich, Axel  
 ; APPLICANT: Gishitsky, Mikhail  
 ; APPLICANT: Sures, Iman G.  
 ; TITLE OF INVENTION: No. 6506578el Megakaryocytic Protein Tyrosine  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232,545  
 ; FILING DATE: 22-APR-1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7683-050  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)790-9741  
 ; TELEFAX: (212)869-9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 16:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 512 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-08-232-545-16  
 Query Match 27.4%; Score 370.5; DB 4; Length 512;  
 Best Local Similarity 39.8%; Pred. No. 1.7e-31;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
 QY 6 SRRKSLPSLSVSSVQGGPVMTMEARSKATAVALGSPAGGPAELSLRLGEPLTIVSED 65  
 DB 38 SNKQRPVE-SQLPGQRFQTKDPEEQGDIVVALYPYDGIHPDDLSPFKGKMKVLEEH 96  
 QY 66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKABELLLPKNPGGAFLIR 121  
 DB 97 GEWMKAKSLTKKEGFIPSNYAKLNTLETEWFFKDIKDAERQQLAPGNSAGAFILR 156

QY 122 ESQTRGSSVLSVRLSRPASWDRIHRHICLDNGWLYISPRLTFFSLQALVDHYSELAD 181  
 DB 157 ESETLKGSSVLSVRDPDPVHGDIKIRSLDNGGYIISPRITPCISDMIKHYKQAD 216  
 QY 182 DICCLKEPCVLRAGPLPK 202  
 DB 217 GLCRLKACI-----SPKPK 233  
 RESULT 3  
 PCT-US95-05008-16  
 ; Sequence 16, Application PC/TUS9505008  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sugen, Inc.  
 ; APPLICANT: 515 Galveston Drive  
 ; APPLICANT: Redwood City, California 94063-4720  
 ; APPLICANT: United States of America  
 ; APPLICANT: Wissenschaften E.V.  
 ; APPLICANT: Hofgarten Str. 2  
 ; APPLICANT: Munchen 80539  
 ; APPLICANT: Germany  
 ; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/05008  
 ; FILING DATE: 24-APR-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/232,545  
 ; FILING DATE: 22-APR-1994  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7683-074  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)790-9090  
 ; TELEFAX: (212)869-9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 16:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 512 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; PCT-US95-05008-16  
 Query Match 27.4%; Score 370.5; DB 5; Length 512;  
 Best Local Similarity 39.8%; Pred. No. 1.7e-31;  
 Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
 QY 6 SRRKSLPSLSVSSVQGGPVMTMEARSKATAVALGSPAGGPAELSLRLGEPLTIVSED 65  
 DB 38 SNKQRPVE-SQLPGQRFQTKDPEEQGDIVVALYPYDGIHPDDLSPFKGKMKVLEEH 96  
 QY 66 GDMWTVLSEVSGREYNIPSVHGVK-----SHGWLYEGLSREKABELLLPKNPGGAFLIR 121  
 DB 97 GEWMKAKSLTKKEGFIPSNYAKLNTLETEWFFKDIKDAERQQLAPGNSAGAFILR 156

QY 122 ESQTRGYSLSVLRSPASWDRIHRYIHCLDNGWLYISPRITFPLSQALVDHYSELAD 181  
 Db 157 ESETLKGSPSLSVDFDVGHDVHKIRSLDNGVYISPRITFPCISDMIKHYKQAD 216  
 QY 182 DICLLKEPCVLQAGPLPK 202  
 Db 217 GLCRRLEKACI----SPKPK 233

RESULT 4  
 US-08-426-509A-17  
 ; Sequence 17, Application US/08426509A  
 ; Patent No. 6326469  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ullrich, Axel  
 ; APPLICANT: Gishizsky, Mikhail  
 ; APPLICANT: Sures, Iman G.  
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
 ; TITLE OF INVENTION: TYROSINE KINASES  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York,  
 ; STATE: NY  
 ; COUNTRY: USA  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/426.509A  
 ; FILING DATE: 21-APR-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/232.545  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7683-0074-999  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 212-790-9090  
 ; TELEFAX: 212-869-9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 17:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 505 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; US-08-426-509A-17

Query Match 26.6%; Score 360.5; DB 4; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30;  
 Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 QY 12 PPSLSVSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGEPPLTVISEDGDWTV 71  
 Db 40 PGPNSHNS---NTPGIREAGSEDIIVVALDYEAITHHEDLSFQKGDQWVLEESGEWKA 96  
 QY 72 LSEVSGREYNIPSVHVGK-----SHGWLYEGLSREKAEELLILPGNPGAFILRESQTRR 127  
 Db 97 RSLATRKEGYIPSNVAVRVDLSLETBEWPFKISRDAERQLLAPGNMLGSPFMRDSEITK 156  
 QY 128 GYSLSVLRSPASWDRIHRYIHCLDNGWLYISPRITFPLSQALVDHYSELADICLL 187  
 Db 157 GYSLSVRDYPDPRQGTGVKHYKIRTLDNGGFIISPRSTFSTLQELVDHYKKGNDGLCQKL 216  
 QY 186 KEPCV 192

RESULT 6  
 PCT-US95-05008-17  
 ; Sequence 17, Application PC/TUS9505008  
 ; GENERAL INFORMATION:

Db 217 SVPCM 221  
 RESULT 5  
 US-08-232-545-17  
 ; Sequence 17, Application US/08232545  
 ; Patent No. 6506578  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ullrich, Axel  
 ; APPLICANT: Gishizsky, Mikhail  
 ; APPLICANT: Sures, Iman G.  
 ; TITLE OF INVENTION: No. 6506578el Megakaryocytic Protein Tyrosine  
 ; TITLE OF INVENTION: Kinases  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232.545  
 ; FILING DATE: 22-APR-1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Coruzzi, Laura A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7683-050  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)790-9090  
 ; TELEFAX: (212)869-9741  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 17:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 505 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-08-232-545-17

Query Match 26.6%; Score 360.5; DB 4; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30;  
 Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 QY 12 PPSLSVSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGEPPLTVISEDGDWTV 71  
 Db 40 PGPNSHNS---NTPGIREAGSEDIIVVALDYEAITHHEDLSFQKGDQWVLEESGEWKA 96  
 QY 72 LSEVSGREYNIPSVHVGK-----SHGWLYEGLSREKAEELLILPGNPGAFILRESQTRR 127  
 Db 97 RSLATRKEGYIPSNVAVRVDLSLETBEWPFKISRDAERQLLAPGNMLGSPFMRDSEITK 156  
 QY 128 GYSLSVLRSPASWDRIHRYIHCLDNGWLYISPRITFPLSQALVDHYSELADICLL 187  
 Db 157 GYSLSVRDYPDPRQGTGVKHYKIRTLDNGGFIISPRSTFSTLQELVDHYKKGNDGLCQKL 216  
 QY 186 KEPCV 192  
 Db 217 SVPCM 221

RESULT 6  
 PCT-US95-05008-17  
 ; Sequence 17, Application PC/TUS9505008  
 ; GENERAL INFORMATION:

APPLICANT: Sugen, Inc.  
 APPLICANT: 515 Galveston Drive  
 APPLICANT: Redwood City, California 94063-4720  
 APPLICANT: United States of America  
 APPLICANT: Wissenschaften E.V.  
 APPLICANT: Hofgarten Str. 2  
 APPLICANT: Munchen 80539  
 APPLICANT: Germany  
 TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
 TITLE OF INVENTION: Kinases  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/05008  
 FILING DATE: 24-APR-1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/232,545  
 FILING DATE: 22-APR-1994  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-074  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212)790-9090  
 TELEFAX: (212)869-9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 505 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 PCT-US95-05008-17

Query Match 26.6%; Score 360.5; DB 5; Length 505;  
 Best Local Similarity 41.6%; Pred. No. 2e-30; Indels 7; Gaps 2;  
 Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
 QY 12 PPSLSVSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPITIVSEDDWTV 71  
 Db 40 PCPNHNS---NTPGIREAGSEDIIVVALXYEAIHEDLSFKQGDQWVLESGEWKA 96  
 QY 72 LSEVSGREYNIPSVHGVK----SHGNLYEGLSRKAEELLPCNPGAPLIRESTOR 127  
 Db 97 RSLATRKGYIPSNVAVRVDLSLETENFPGISRKDAERQLLAPNMLGSMFMRDSEITK 156  
 QY 128 GYSLSVLSRSPASWDRIHRIHCLDNGWLYISRLTFPSLQALVDHYSELADICLL 187  
 Db 157 GYSLSVLDYDPQGTGVKHYKIRTLDNGGFVSPRSTFTLQELVDHYKKGNDGLCQKL 216  
 QY 188 KEPCV 192  
 Db 217 SVPCM 221

RESULT 7  
 US-08-426-509A-19  
 ; Sequence 19, Application US/08426509A  
 ; Patent No. 6326469

GENERAL INFORMATION:  
 APPLICANT: Ullrich, Axel  
 APPLICANT: Gishizky, Mikhail  
 APPLICANT: Sures, Irman G.  
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
 TITLE OF INVENTION: TYROSINE KINASES  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds  
 STREET: 1155 Avenue of the Americas  
 CITY: New York,  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/426,509A  
 FILING DATE: 21-APR-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/232,545  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-0074-999  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-790-9090  
 TELEFAX: 212-869-9741  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 499 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: No. 6326469e  
 US-08-426-509A-19  
 Query Match 25.5%; Score 344.5; DB 4; Length 499;  
 Best Local Similarity 36.0%; Pred. No. 1.1e-28;  
 Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;  
 QY 1 MGSLSFRKSLPSLSVSSVQGGPV-----TMEARSK 34  
 Db 1 MGLLSKRGQ-----VSEKKGWSPVKIRTDKAPPLPPLVFNHAPPSPNQDPDEEB 54  
 QY 35 ATAVALGSPFAGGPAELSLRLGEPITIVSEDDWTVLSEVSGREYNIPSVHGVKVS--- 91  
 Db 55 RFVVALFDYAAVNDRLDQVKGKQLVLRSTGDWMLARSLVTGREGYVPSNFVAPVETLE 114  
 QY 92 -HGWLVEGLSRKAEELLPCNPGAPLIRESTORRGSYSLSVLSRSPASWDRIHRI 150  
 Db 115 VEKWFRTISRDAERQLLAPNMGKAGSLIRESESNKGAFLSVK-DITTOGEVVKHYKI 173  
 QY 151 HCLDNGWLYISRLTFPSLQALVDHYSELADICLLKEPCV 192  
 Db 174 RSLDNGGYISPRITFTLQALVQHYSKRGDGLCQKLTLPCV 215  
 RESULT 8  
 US-08-232-545-19  
 ; Sequence 19, Application US/08232545  
 ; Patent No. 6506578  
 ; GENERAL INFORMATION:  
 APPLICANT: Ullrich, Axel  
 APPLICANT: Gishizky, Mikhail  
 APPLICANT: Sures, Irman G.  
 TITLE OF INVENTION: No. 6506578e1 Megakaryocytic Protein Tyrosine

TITLE OF INVENTION: Kinases  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,545  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-050  
TELEPHONE: (212)790-9090  
TELEFAX: (212)869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 499 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US-08-232-545-19

Query Match 25.5%; Score 344.5; DB 4; Length 499;

Best Local Similarity 36.0%; Pred. No. 1.1e-28;  
Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;

QY 1 MGLSPRRKSLPSPSSSVQGGPV-----TWAEERSK 34  
DB 1 MGLLSKRG-----VSEKKGWSPVKIRTDQKAPPLPLVFNHAPPSPNQPDDEE 54  
QY 35 ATAVALGSPAGPAELSLRLGEPLTVISGDWTVLSEVSGREYNIPSVHVGVKVS--- 91  
DB 55 RFVVALFDYAANDRDLOVLKGEKQLVLRSTGDWMLARSLVTGREGVPSNFPVAPVETLE 114  
QY 92 -HGWLVEGLSREKAEELLPLPGNPGGAFLIRESQTRGYSLSVLSRPPASWDRIHYRI 150  
DB 115 VEKWFRTISRKDAERQLLAPNKGAGFLIRESSEKNGAFSLSVK-DITTOGEVVVKHYKI 173  
QY 151 HCLDNGWLYISPLTPPSLQALVDHYSELADDCILLKEPCV 192  
DB 174 RSLDNGGYIISPRITFTLQALVQHYSKKGDLGCLQKLTLPV 215

## RESULT 9

PCT-US95-05008-19  
Sequence 19, Application PC/TUS9505008  
GENERAL INFORMATION:  
APPLICANT: Sugen, Inc.  
APPLICANT: 515 Galveston Drive  
APPLICANT: Redwood City, California 94063-4720  
APPLICANT: United States of America  
APPLICANT: Wissenschaften E.V.  
APPLICANT: Hofgarten Str. 2  
APPLICANT: Munchen 80539  
APPLICANT: Germany  
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05008  
FILING DATE: 24-APR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/232,545  
FILING DATE: 22-APR-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-074  
TELEPHONE: (212)790-9090  
TELEFAX: (212)869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 499 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US95-05008-19

Query Match 25.5%; Score 344.5; DB 5; Length 499;

Best Local Similarity 36.0%; Pred. No. 1.1e-28;  
Matches 80; Conservative 32; Mismatches 73; Indels 37; Gaps 4;

QY 1 MGLSPRRKSLPSPSSSVQGGPV-----TWAEERSK 34  
DB 1 MGLLSKRG-----VSEKKGWSPVKIRTDQKAPPLPLVFNHAPPSPNQPDDEE 54  
QY 35 ATAVALGSPAGPAELSLRLGEPLTVISGDWTVLSEVSGREYNIPSVHVGVKVS--- 91  
DB 55 RFVVALFDYAANDRDLOVLKGEKQLVLRSTGDWMLARSLVTGREGVPSNFPVAPVETLE 114  
QY 92 -HGWLVEGLSREKAEELLPLPGNPGGAFLIRESQTRGYSLSVLSRPPASWDRIHYRI 150  
DB 115 VEKWFRTISRKDAERQLLAPNKGAGFLIRESSEKNGAFSLSVK-DITTOGEVVVKHYKI 173  
QY 151 HCLDNGWLYISPLTPPSLQALVDHYSELADDCILLKEPCV 192  
DB 174 RSLDNGGYIISPRITFTLQALVQHYSKKGDLGCLQKLTLPV 215

## RESULT 10

US-09-862-154-1  
Sequence 1, Application US/09862154  
Patent No. 6589758  
GENERAL INFORMATION:  
APPLICANT: Zhu, Xiaocian  
TITLE OF INVENTION: Crystal of a Kinase-ligand Complex and Methods of Use  
FILE REFERENCE: Atty. Docket No. 6589758; A-749  
CURRENT APPLICATION NUMBER: US/09/862,154  
CURRENT FILING DATE: 2001-05-21  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 1  
LENGTH: 508  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-862-154-1

Query Match 25.1%; Score 340; DB 4; Length 508;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

QY 26 VTMEARSKAT-----AVALGSPFAGGPAELSLRLGELPLTIVSDGDWMTVLSEVSGRE 79  
DB 48 VTGEGNPPASPLODNLVIALHSEYPSHDGDLGFEKGEQLRIEQSGEWWKQAQSLTTGQE 107  
QY 80 YNIPSVHVGVKVS---HGMWLYEGLSREKAEELLLPQNGGAFILRESOTRRGYSLSVR 135  
DB 108 GFIPFNFVAKANLEPEPWFKNLSRDAERQLAPGNTHGSLFIRESESTAGSFSLSVR 167  
QY 136 LSRPASWDRIHRYHICLDNGWLYISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 195  
DB 168 DFDQNGQEVVVKYKIRNLNDNGGFIISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 227

## RESULT 11

US-09-039-555B-17  
; Sequence 17, Application US/09039555B  
; Patent No. 6033856  
; GENERAL INFORMATION:  
; APPLICANT: Koerner, Kathrin  
; APPLICANT: Mueller, Rolf  
; APPLICANT: Sadlacek, Hans-Harald  
; TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS  
; PREPARATION AND USE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,555B  
; FILING DATE: 16-MAR-1998  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 19710643.9  
; FILING DATE: 14-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 016779/0131  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 509 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: linear  
; MOLECULE TYPE: protein  
US-09-039-555B-17

Query Match 25.1%; Score 340; DB 3; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

QY 26 VTMEARSKAT-----AVALGSPFAGGPAELSLRLGELPLTIVSDGDWMTVLSEVSGRE 79  
DB 49 VTGEGNPPASPLODNLVIALHSEYPSHDGDLGFEKGEQLRIEQSGEWWKQAQSLTTGQE 108

QY 80 YNIPSVHVGVKVS---HGMWLYEGLSREKAEELLLPQNGGAFILRESOTRRGYSLSVR 135  
DB 109 GFIPFNFVAKANLEPEPWFKNLSRDAERQLAPGNTHGSLFIRESESTAGSFSLSVR 168  
QY 136 LSRPASWDRIHRYHICLDNGWLYISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 195  
DB 169 DFDQNGQEVVVKYKIRNLNDNGGFIISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 228

## RESULT 12

US-08-426-509A-18  
; Sequence 18, Application US/08426509A  
; Patent No. 6326469  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Gishizky, Mikhail  
; APPLICANT: Suresh, Iman G.  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN  
; TYROSINE KINASES  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York,  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/426,509A  
; FILING DATE: 21-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/232,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-0074-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-869-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 509 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
US-08-426-509A-18

Query Match 25.1%; Score 340; DB 4; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

QY 26 VTMEARSKAT-----AVALGSPFAGGPAELSLRLGELPLTIVSDGDWMTVLSEVSGRE 79  
DB 49 VTGEGNPPASPLODNLVIALHSEYPSHDGDLGFEKGEQLRIEQSGEWWKQAQSLTTGQE 108  
QY 80 YNIPSVHVGVKVS---HGMWLYEGLSREKAEELLLPQNGGAFILRESOTRRGYSLSVR 135  
DB 109 GFIPFNFVAKANLEPEPWFKNLSRDAERQLAPGNTHGSLFIRESESTAGSFSLSVR 168  
QY 136 LSRPASWDRIHRYHICLDNGWLYISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 195  
DB 169 DFDQNGQEVVVKYKIRNLNDNGGFIISPRITFPGLHVLVRYHYNASDGLCTRLSRPCQTOX 228

## RESULT 13

US-09-457-040B-8  
; Sequence 8, Application US/09457040B  
; Patent No. 6387641  
; GENERAL INFORMATION:  
; APPLICANT: Vertex Pharmaceuticals Incorporated  
; APPLICANT: Bellon, Steve  
; TITLE OF INVENTION: Crystallized P38 Complexes  
; FILE REFERENCE: VPI/98-14  
; CURRENT APPLICATION NUMBER: US/09/457,040B  
; CURRENT FILING DATE: 1999-12-08  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Human  
US-09-457-040B-8

Query Match 25.1%; Score 340; DB 4; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
  
QY 26 VTMEARSKAT-----AVALGSPAGGPAELSLRLGPEPLTIVSEGDGWTIVLSEVSGRE 79  
Db 49 VTVEGSPASPLQDNLVIALHSYEPHSDGDLGFEKGEQLRILEQSGEWKQAQLTTGQE 108  
  
QY 80 YNIPSVHVGVKVS-----HGMVYEGLSREKAEELLILPGNPGGAFILRESQTRRGSYSLSVR 135  
Db 109 GTIFNFVAKANSLEPEPFKNLSRKAERQLLAPGNTHGSLFIRESESTAGSFSLSVR 169  
  
QY 136 LSRPASWDRIHYRIHCLDNGWLYISPRITFFPSLQALVDHYSELADDCILLKEPCVLQR 195  
Db 169 DFDQNGQEVVKYKIRNLNDGGFYISPRITFFGLHVLVRYHTNASDGLCTRLSRPCQTK 228

RESULT 14  
US-08-232-545-18  
; Sequence 18, Application US/08232545  
; Patent No. 6506578  
; GENERAL INFORMATION:  
; APPLICANT: Ulrich, Axel  
; APPLICANT: Gishizky, Mikhail  
; APPLICANT: Sures, Irmann G.  
; TITLE OF INVENTION: Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,545  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-050  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212)869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 509 amino acids

; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-232-545-18  
  
Query Match 25.1%; Score 340; DB 4; Length 509;  
Best Local Similarity 40.6%; Pred. No. 3.5e-28;  
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;  
  
QY 26 VTMEARSKAT-----AVALGSPAGGPAELSLRLGPEPLTIVSEGDGWTIVLSEVSGRE 79  
Db 49 VTVEGSPASPLQDNLVIALHSYEPHSDGDLGFEKGEQLRILEQSGEWKQAQLTTGQE 108  
  
QY 80 YNIPSVHVGVKVS-----HGMVYEGLSREKAEELLILPGNPGGAFILRESQTRRGSYSLSVR 135  
Db 109 GTIFNFVAKANSLEPEPFKNLSRKAERQLLAPGNTHGSLFIRESESTAGSFSLSVR 169  
  
QY 136 LSRPASWDRIHYRIHCLDNGWLYISPRITFFPSLQALVDHYSELADDCILLKEPCVLQR 195  
Db 169 DFDQNGQEVVKYKIRNLNDGGFYISPRITFFGLHVLVRYHTNASDGLCTRLSRPCQTK 228

RESULT 15  
PCT-US95-05008-18  
; Sequence 18, Application PC/TUS9505008  
; GENERAL INFORMATION:  
; APPLICANT: Sugen, Inc.  
; APPLICANT: 515 Galveston Drive  
; APPLICANT: Redwood City, California 94063-4720  
; APPLICANT: United States of America  
; APPLICANT: Wissenschaften E.V.  
; APPLICANT: Hofgarten Str. 2  
; APPLICANT: Munchen 80539  
; APPLICANT: Germany  
; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
; TITLE OF INVENTION: Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05008  
; FILING DATE: 24-APR-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/232,545  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-074  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)790-9090  
; TELEFAX: (212)869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 509 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein

PCT-US95-05008-18

```
Query Match      25.1%; Score 340; DB 5; Length 509;
Best Local Similarity 40.6%; Pred. No. 3.5e-28;
Matches 73; Conservative 26; Mismatches 71; Indels 10; Gaps 2;

Qy 26 VTWEAERSKAT-----AVALGSPAGGPAELSLRLGELTIVSDDGDMWTVLSEVSGRE 79
Db 49 VTVEGSPNPPASPLQDNLVIALHSYEPHSDGLGFEKGEQLRILEQSGEWWKQSLTTGQE 108

Qy 80 YNIPSVHVGVKS----HGWLYEGLSRKAEELLFPNPGCAFIREQTRRGSYSLSVR 135
Db 109 GFIPFNFVAXANSLEPEPWFKKLSRKDAERQLLAPGNTHGSLFIRESESTAGSFSLSVR 168

Qy 136 LSRPASWDRIRHYRIHCLDNGWLYISRLTFPSLQALVDHYSELADDICLLKEPCVLQR 195
Db 169 DFDQNGQEVVVKHYKIRNLNNGGFYISPRITFPGLHELVRHYTNASDGLCTRUSRPCQOK 228
```

Search completed: February 20, 2004, 09:34:03  
Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:32:17 ; Search time 593 Seconds  
(without alignments)  
92.936 Million cell updates/sec

Title: US-09-939-853A-75  
Perfect score: 1353  
Sequence: 1 MGSUFSRKSLSPLSSSV.....RESLSFVLSNDEAVSLDDA 261

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1347	99.6	261	14	US-10-043-649-2
2	826	61.0	159	9	US-09-867-550-954
3	586	43.3	113	9	US-09-867-550-1916
4	481.5	35.6	276	9	US-09-870-759-64
5	481.5	35.6	276	10	US-09-751-708A-64
6	481.5	35.6	276	14	US-10-043-649-3
7	452.5	33.4	96	9	US-09-867-550-952
8	370.5	27.4	511	15	US-10-394-322A-42
9	370.5	27.4	512	9	US-09-977-269-16
10	370.5	27.4	512	9	US-09-977-269-16
11	370.5	27.4	512	10	US-09-977-261-16
12	370.5	27.4	512	15	US-10-116-275-162
13	360.5	26.6	505	9	US-09-977-269-17
14	360.5	26.6	505	9	US-09-977-260-17
15	360.5	26.6	505	10	US-09-977-261-17

16	360.5	26.6	505	15	US-10-193-720-2	Sequence 2, Appli
17	360.5	26.6	526	15	US-10-394-322A-31	Sequence 31, Appli
18	352.5	26.1	504	15	US-10-394-322A-4	Sequence 4, Appli
19	352.5	26.1	505	9	US-09-771-161A-186	Sequence 186, Appl
20	348	25.7	509	15	US-10-366-288-28	Sequence 28, Appl
21	344.5	25.5	499	9	US-09-977-269-19	Sequence 19, Appl
22	344.5	25.5	499	9	US-09-977-260-19	Sequence 19, Appl
23	344.5	25.5	499	10	US-09-977-261-19	Sequence 19, Appl
24	340	25.1	508	15	US-10-394-322A-41	Sequence 41, Appl
25	340	25.1	509	9	US-09-977-269-18	Sequence 18, Appl
26	340	25.1	509	9	US-09-977-260-18	Sequence 18, Appl
27	340	25.1	509	10	US-09-977-261-18	Sequence 18, Appl
28	340	25.1	509	14	US-10-212-346-1	Sequence 1, Appli
29	322.5	23.8	454	9	US-09-771-161A-95	Sequence 95, Appl
30	318.5	23.5	537	9	US-09-771-161A-212	Sequence 212, App
31	318.5	23.5	537	9	US-09-771-161A-213	Sequence 213, App
32	315.5	23.3	311	9	US-09-771-161A-121	Sequence 121, App
33	315.5	23.3	387	9	US-09-771-161A-122	Sequence 122, App
34	315.5	23.3	536	15	US-10-394-322A-30	Sequence 30, Appl
35	315.5	23.3	537	9	US-09-977-269-11	Sequence 11, Appl
36	315.5	23.3	537	9	US-09-977-260-11	Sequence 11, Appl
37	315.5	23.3	537	10	US-09-977-261-11	Sequence 11, Appl
38	315.5	23.3	543	9	US-09-977-269-14	Sequence 14, Appl
39	315.5	23.3	543	9	US-09-977-260-14	Sequence 14, Appl
40	315.5	23.3	543	10	US-09-977-261-14	Sequence 14, Appl
41	315.5	23.3	543	14	US-10-298-377A-4	Sequence 4, Appli
42	315.5	23.3	543	15	US-10-394-322A-68	Sequence 68, Appl
43	312.5	23.1	529	9	US-09-977-269-15	Sequence 15, Appl
44	312.5	23.1	529	9	US-09-977-260-15	Sequence 15, Appl
45	312.5	23.1	529	10	US-09-977-261-15	Sequence 15, Appl

## ALIGNMENTS

### RESULT 1

US-10-043-649-2  
; Sequence 2, Application US/10043649  
; Publication No. US20030059924A1  
; GENERAL INFORMATION:  
; APPLICANT: Hollander, Sacha J.  
; APPLICANT: Mendenhall, Marcy K.  
; APPLICANT: Pardo, Jorge  
; APPLICANT: Spencer, Collin  
; APPLICANT: Fu, C. Alan  
; APPLICANT: Luo, Ying  
; APPLICANT: Payan, Donald G.  
; APPLICANT: Mancebo, Helena S.Y.  
; APPLICANT: Wu, Jun  
; APPLICANT: Zhou, Xiulan  
; APPLICANT: Shen, Mary  
; APPLICANT: Liao, X. Charlene  
; APPLICANT: Sheng, Ning  
; TITLE OF INVENTION: Cloning of a No. US20030059924A1e1 Inhibitor of Antigen-receptor ;  
; FILE REFERENCE: A-70219-1/RMS/DHR  
; CURRENT APPLICATION NUMBER: US/10/043,649  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/260,953  
; PRIOR FILING DATE: 2001-01-10  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-043-649-2

Query Match 99.6%; Score 1347; DB 14; Length 261;  
Best Local Similarity 99.6%; Pred. No. 2e-124;  
Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSUFSRKSLSPLSSSVQGGQPTMEARSKATAVALGSPAGPAAELSLRLGEPILT 60

Db 1 MGSLPSRRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPLT 60  
Qy 61 IVSEDCDMMWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLPGNPGGAFLI 120  
Db 61 IVSEDCDMMWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLPGNPGGAFLI 120  
Qy 121 RESQTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLVIISPRLTTPPSIQALVDHYSELA 180  
Db 121 RESQTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLVIISPRLTTPPSIQALVDHYSELA 180  
Qy 181 DDICCLLKPCVQLORAGPLGKDIPLPVTVORTPLANKKELSSLLFSEATGEESLLSG 240  
Db 181 DDICCLLKPCVQLORAGPLGKDIPLPVTVORTPLANKKELSSLLFSEATGEESLLSG 240  
Qy 241 LRESLSFYISLNDEAVSLDDA 261  
Db 241 LRESLSFYISLNDEAVSLDDA 261

RESULT 2  
US-09-867-550-954  
; Sequence 954, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fud, Pamela  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 954  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-867-550-954

Query Match 61.0%; Score 826; DB 9; Length 159;  
Best Local Similarity 99.4%; Pred. No. 2.3e-73;  
Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MGSLPSRRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPLT 60  
Db 1 MGSLPSRRKSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPLT 60  
Qy 61 IVSEDCDMMWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLPGNPGGAFLI 120  
Db 61 IVSEDCDMMWTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLPGNPGGAFLI 120  
Qy 121 RESQTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLVIISPRLTTPPSIQALVDHYSELA 159  
Db 121 RESQTRRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLVIISPRLTTPPSIQALVDHYSELA 159

RESULT 3  
US-09-867-550-1916  
; Sequence 1916, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fud, Pamela  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and

; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1916  
; LENGTH: 113  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)  
; OTHER INFORMATION: Wherein Xaa may be any one of Arg or Gly or Trp  
US-09-867-550-1916

Query Match 43.3%; Score 586; DB 9; Length 113;  
Best Local Similarity 100.0%; Pred. No. 6.8e-50;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 150 IHCLDNGWLVIISPRLTTPPSIQALVDHYSELAADDICCLLKPCVQLORAGPLGKDIPLPVT 209  
Db 2 IHCLDNGWLVIISPRLTTPPSIQALVDHYSELAADDICCLLKPCVQLORAGPLGKDIPLPVT 61  
Qy 210 VORTPLANKKELSSLLFSEATGEESLLSGLRESLSFYISLNDEAVSLDDA 261  
Db 62 VORTPLANKKELSSLLFSEATGEESLLSGLRESLSFYISLNDEAVSLDDA 113

RESULT 4  
US-09-870-759-64  
; Sequence 64, Application US/09870759  
; Patent No. US20020177551A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; CURRENT FILING DATE: 2002-01-14  
; PRIOR APPLICATION NUMBER: US 60/208,128  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 64  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-759-64

Query Match 35.6%; Score 481.5; DB 9; Length 276;  
Best Local Similarity 40.3%; Pred. No. 4.8e-39;  
Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;  
Qy 9 KSLPSPSLSSVQGGPVTMEARSKATAVALGSPFAGGPAELSLRLGEPLTIVSEDCGW 68  
Db 6 KSTPAPA-----ERPLNPEGLSDDFLAVLSDFSPDISPPIFRGKELRVISDEGGM 58  
Qy 69 WTVLSEVSGREYNIPSVHVKVSHGWLVEGLSREKAEELLLPGNPGGAFLIRESQTRRG 128  
Db 59 WKALISLTGRESYIPGICVARVTHGWFEGGLORDKAEELLQLPDTKVGSMIRSEETKKG 118  
Qy 129 SYSLSVRLSRPASWDRIHYRIHCLDNGWLVIISPRLTTPPSIQALVDHYSELAADDICCLLK 188  
Db 119 FYSLSVR-----HRQVKHYRIFRLPNWYIISPRLTFCLEDLVNHYSEVADGLCCVLT 172  
Qy 189 EPCVQLORAGPLGKDIPLPVTVORTPLANKKELSSLLFSEATG-----EESLLSEGL 241  
Db 173 TPCLTQSTAAPAVRASSSPVTLRQKTVDMWRVSR---LQEDPEGTENPLGVDESLSFSG 229  
Qy 242 RESLSFYISLNDE 254

```

Db      230 RESIASYLSLTSE 242

RESULT 5
US-09-751-708A-64
; Sequence 64, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751.708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-64

Query Match      35.6%; Score 481.5; DB 10; Length 276;
Best Local Similarity 40.3%; Pred. No. 4.8e-39;
Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;

QY      9 KSLPSPSSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGEPPLTIVSEGDW 68
Db      6 KSTPAPA-----ERPLNPEGLSDDFLAVLSVPSDPISPPFRGEGKLRVSDSGW 58

QY      69 WTVLSEVSGREYNIPSVHGVKSHGWLYEGLSREKAEELLPLGNPGGAFLIRESQTRG 128
Db      59 WKALSLSGTRESYIPGICVARVYHGWLFEGLDKAEELLQLPDTKVSFMRSEETKKG 118

QY      129 SYSLSVRLSRPASWDRIHYRHICLDNGWLYTSPRLTTPSLQALVDHYSELADDICLLK 188
Db      119 FYLSVR-----HRQVKHYRIFRLPNWYIISPLTTCQLEDLVNHYSEVADGLCCVLT 172

QY      189 EPCVLQAGPLPKDIPVTVTQRTPLNWKELDSSLFSEATG-----EESLLSEGL 241
Db      173 TPCLTQSTAAPAVRASSSPVTLRQKTVDWRRVSR---LQEDPEGTENPLGVDESLSFYGL 229

QY      242 RESLSFYISLND 254
Db      230 RESIASYLSLTSE 242

RESULT 6
US-10-043-649-3
; Sequence 3, Application US/10043649
; Publication No. US20030059924A1
; GENERAL INFORMATION:
; APPLICANT: Holland, Sacha J.
; APPLICANT: Mendenhall, Marcy K.
; APPLICANT: Pardo, Jorge
; APPLICANT: Spencer, Collin
; APPLICANT: Fu, C. Alan
; APPLICANT: Luo, Ying
; APPLICANT: Payan, Donald G.
; APPLICANT: Mancebo, Helena S.Y.
; APPLICANT: Wu, Jun
; APPLICANT: Zhou, Xiulan
; APPLICANT: Shen, Mary
; APPLICANT: Liao, X. Charlene
; APPLICANT: Sheng, Ning
; TITLE OF INVENTION: Cloning of a No. US20030059924A1a1 Inhibitor of Antigen-receptor
; FILE REFERENCE: A-70219-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/10/043,649
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/260,953
; PRIOR FILING DATE: 2001-01-10

US-09-939-853a-75.rapb
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-043-649-3

Query Match      35.6%; Score 481.5; DB 14; Length 276;
Best Local Similarity 40.3%; Pred. No. 4.8e-39;
Matches 102; Conservative 43; Mismatches 85; Indels 23; Gaps 4;

QY      9 KSLPSPSSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGEPPLTIVSEGDW 68
Db      6 KSTPAPA-----ERPLNPEGLSDDFLAVLSVPSDPISPPFRGEGKLRVSDSGW 58

QY      69 WTVLSEVSGREYNIPSVHGVKSHGWLYEGLSREKAEELLPLGNPGGAFLIRESQTRG 128
Db      59 WKALSLSGTRESYIPGICVARVYHGWLFEGLDKAEELLQLPDTKVSFMRSEETKKG 118

QY      129 SYSLSVRLSRPASWDRIHYRHICLDNGWLYTSPRLTTPSLQALVDHYSELADDICLLK 188
Db      119 FYLSVR-----HRQVKHYRIFRLPNWYIISPLTTCQLEDLVNHYSEVADGLCCVLT 172

QY      189 EPCVLQAGPLPKDIPVTVTQRTPLNWKELDSSLFSEATG-----EESLLSEGL 241
Db      173 TPCLTQSTAAPAVRASSSPVTLRQKTVDWRRVSR---LQEDPEGTENPLGVDESLSFYGL 229

QY      242 RESLSFYISLND 254
Db      230 RESIASYLSLTSE 242

RESULT 7
US-09-867-550-952
; Sequence 952, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1a1 Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 952
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-952

Query Match      33.4%; Score 452.5; DB 9; Length 96;
Best Local Similarity 76.8%; Pred. No. 7.9e-37;
Matches 96; Conservative 0; Mismatches 0; Indels 29; Gaps 1;

QY      1 MGSLSRHRKSLPSPSSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGEPPLT 60
Db      1 MGSLSRHRKSLPSPSSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGEPPLT 60

QY      61 IVSEGDGDMWTVLSEVSGREYNIPSVHGVKSHGWLYEGLSREKAEELLPLGNPGGAFLI 120
Db      61 IVSE-----WLVEGLSREKAEELLPLGNPGGAFLI 91

QY      121 RESQT 125
Db      92 RESQT 96

```

RESULT 8  
US-10-394-322A-42  
; Sequence 42, Application US/10394322A  
; Publication No. US20030232391A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNEIS PHARMACEUTICALS, INC.  
; APPLICANT: Prescott, John C.  
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
; FILE REFERENCE: 39750-0006 US  
; CURRENT APPLICATION NUMBER: US/10/394,322A  
; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: US 60/366,892  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42  
; LENGTH: 511  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-394-322A-42

Query Match 27.4%; Score 370.5; DB 15; Length 511;  
Best Local Similarity 39.8%; Pred. No. 1e-27;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
  
QY 6 SRRKSLPSPSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGELPTIVSED 65  
DB 37 SNKQRPVPE-SQLLPQRFQTKDPEEQDIVVALYPYDGIHPDDLSFKKGKMKVLEH 95  
  
QY 66 GDWWTVLSEVSGREYNIPSPVHVGVK-----SHGWLYEGLSREKAEELLLLPGNPGGAFLIR 121  
DB 96 GEWWAKASLLTKKEGFIPSNYVAKLNTLETEWFFKDIITRKDAERQLLAFGNSAGAFLIR 155  
  
QY 122 ESQTRRGYSLSVRLSRPASWDRIHRYIHCLDNGWLYISPRLTTPPSLQALVDHYSELAD 181  
DB 156 ESETLKGFSLSVRDFFPVHGDVIRKIRSLDNGGYIISPRITFPFCISDMIKHYQKQAD 215  
  
QY 182 DICLLKEPCVLQAGPLPGK 202  
DB 216 GLCRRLEKACI-----SPKPK 232

RESULT 9  
US-09-977-269-16  
; Sequence 16, Application US/09977269  
; Patent No. US20020082037A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1260  
; CURRENT APPLICATION NUMBER: US/09/977,269  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-269-16

Query Match 27.4%; Score 370.5; DB 9; Length 512;  
Best Local Similarity 39.8%; Pred. No. 1e-27;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
  
QY 6 SRRKSLPSPSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGELPTIVSED 65  
DB 38 SNKQRPVPE-SQLLPQRFQTKDPEEQDIVVALYPYDGIHPDDLSFKKGKMKVLEH 96

QY 66 GDWWTVLSEVSGREYNIPSPVHVGVK-----SHGWLYEGLSREKAEELLLLPGNPGGAFLIR 121  
DB 97 GEWWAKASLLTKKEGFIPSNYVAKLNTLETEWFFKDIITRKDAERQLLAFGNSAGAFLIR 156  
  
QY 122 ESQTRRGYSLSVRLSRPASWDRIHRYIHCLDNGWLYISPRLTTPPSLQALVDHYSELAD 181  
DB 157 ESETLKGFSLSVRDFFPVHGDVIRKIRSLDNGGYIISPRITFPFCISDMIKHYQKQAD 216  
  
QY 182 DICLLKEPCVLQAGPLPGK 202  
DB 217 GLCRRLEKACI-----SPKPK 233

RESULT 10  
US-09-977-260-16  
; Sequence 16, Application US/09977260  
; Publication No. US20020192790A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1260  
; CURRENT APPLICATION NUMBER: US/09/977,260  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 512  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-260-16

Query Match 27.4%; Score 370.5; DB 9; Length 512;  
Best Local Similarity 39.8%; Pred. No. 1e-27;  
Matches 80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;  
  
QY 6 SRRKSLPSPSSSVQGGPVTMEARSKATAVALGSPAGPAELSLRLGELPTIVSED 65  
DB 38 SNKQRPVPE-SQLLPQRFQTKDPEEQDIVVALYPYDGIHPDDLSFKKGKMKVLEH 96  
  
QY 66 GDWWTVLSEVSGREYNIPSPVHVGVK-----SHGWLYEGLSREKAEELLLLPGNPGGAFLIR 121  
DB 97 GEWWAKASLLTKKEGFIPSNYVAKLNTLETEWFFKDIITRKDAERQLLAFGNSAGAFLIR 156  
  
QY 122 ESQTRRGYSLSVRLSRPASWDRIHRYIHCLDNGWLYISPRLTTPPSLQALVDHYSELAD 181  
DB 157 ESETLKGFSLSVRDFFPVHGDVIRKIRSLDNGGYIISPRITFPFCISDMIKHYQKQAD 216  
  
QY 182 DICLLKEPCVLQAGPLPGK 202  
DB 217 GLCRRLEKACI-----SPKPK 233

RESULT 11  
US-09-977-261-16  
; Sequence 16, Application US/09977261  
; Publication No. US20030054527A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1259  
; CURRENT APPLICATION NUMBER: US/09/977,261  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 16
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-261-16

Query Match      27.4%; Score 370.5; DB 10; Length 512;
Best Local Similarity 39.8%; Pred. No. 1e-27;
Matches      80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY  6 SRKSLPSPSSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSD 65
DB  38 SNKQRPVPE-SQLLPGQRFQTKPEEGDIVVLYPDGHPDDLSPFKGKMKVLEH 96
QY  66 GDMWTVLSEVSGREYNIPSVHVGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIR 121
DB  97 GEMWAKSLITKKEGFIPTNVAKLNTLEETEEWFFKDIITKDAERQLLAPGNSAGAF 156
QY  122 ESQTRGYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRITPPSLQALVDHYSELAD 181
DB  157 ESETLKGFSLSVRDFDPVHGVDVVKHYKIRSLDNGGYISPRITPPCISDMIKHYQKQAD 216
QY  182 DICLLKEPCVLQAGPLPGK 202
DB  217 GLCRLEKACI-----SPKPK 233

RESULT 12
US-10-116-275-162
; Sequence 162, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elian Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 162
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-162

Query Match      27.4%; Score 370.5; DB 15; Length 512;
Best Local Similarity 39.8%; Pred. No. 1e-27;
Matches      80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY  6 SRKSLPSPSSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSD 65
DB  38 SNKQRPVPE-SQLLPGQRFQTKPEEGDIVVLYPDGHPDDLSPFKGKMKVLEH 96
QY  66 GDMWTVLSEVSGREYNIPSVHVGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIR 121
DB  97 GEMWAKSLITKKEGFIPTNVAKLNTLEETEEWFFKDIITKDAERQLLAPGNSAGAF 156
QY  122 ESQTRGYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRITPPSLQALVDHYSELAD 181
DB  157 ESETLKGFSLSVRDFDPVHGVDVVKHYKIRSLDNGGYISPRITPPCISDMIKHYQKQAD 216
QY  182 DICLLKEPCVLQAGPLPGK 202
DB  217 GLCRLEKACI-----SPKPK 233

RESULT 13
US-10-116-275-162
; Sequence 162, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elian Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 162
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-162

Query Match      27.4%; Score 370.5; DB 15; Length 512;
Best Local Similarity 39.8%; Pred. No. 1e-27;
Matches      80; Conservative 36; Mismatches 76; Indels 9; Gaps 3;

QY  6 SRKSLPSPSSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSD 65
DB  38 SNKQRPVPE-SQLLPGQRFQTKPEEGDIVVLYPDGHPDDLSPFKGKMKVLEH 96
QY  66 GDMWTVLSEVSGREYNIPSVHVGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIR 121
DB  97 GEMWAKSLITKKEGFIPTNVAKLNTLEETEEWFFKDIITKDAERQLLAPGNSAGAF 156
QY  122 ESQTRGYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRITPPSLQALVDHYSELAD 181
DB  157 ESETLKGFSLSVRDFDPVHGVDVVKHYKIRSLDNGGYISPRITPPCISDMIKHYQKQAD 216
QY  182 DICLLKEPCVLQAGPLPGK 202
DB  217 GLCRLEKACI-----SPKPK 233
```

```
US-09-977-269-17
; Sequence 17, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,269
; CURRENT FILING DATE: 2001-10-16
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-269-17

Query Match      26.6%; Score 360.5; DB 9; Length 505;
Best Local Similarity 41.6%; Pred. No. 9.6e-27;
Matches      77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;

QY  12 PPSLSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSDGDMWTV 71
DB  40 PGNSHNS---NTPGIREAGSEDIIVVLYDYEAHHEHEDLSFKQGDQMVVLEESGEWKA 96
QY  72 LSEVSGREYNIPSVHVGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIRESQTRR 127
DB  97 RSLATRKEGYPISNYVAVRVDLSLEETEEWFFKGISRDAERQLLAPGNNLGSFMIRDSEITK 156
QY  128 GYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRITPPSLQALVDHYSELADICCLL 187
DB  157 GYSLSVRYDPRQGVTVKHYKIRTLNDNGGFYISPRSTETLQELVDHYKKGNDGLCQKL 216
QY  188 KEPCV 192
DB  217 SVPCM 221

RESULT 14
US-09-977-260-17
; Sequence 17, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; CURRENT FILING DATE: 2001-10-16
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-260-17

Query Match      26.6%; Score 360.5; DB 9; Length 505;
Best Local Similarity 41.6%; Pred. No. 9.6e-27;
Matches      77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;

QY  12 PPSLSVQGGPVTMEASRSKATAVALGSPAGGPAELSLRLGELPTIVSDGDMWTV 71
DB  40 PGNSHNS---NTPGIREAGSEDIIVVLYDYEAHHEHEDLSFKQGDQMVVLEESGEWKA 96
QY  72 LSEVSGREYNIPSVHVGVK-----SHGWLYEGLSREKAEELLPLPGNPGGAFLIRESQTRR 127
```

Db 97 RSLATKEGYIPSNYVARVDSLETEEFKGISRKAERQLLAPGNMLGSMFMRDSEYTK 156  
Qy 128 GSYSLVSLSRPASWDRIRHYRHHCLDNGWLYISPLTFPSLQALVDHYSELADDCILL 187  
Db 157 GSYSLVSRVDYDPRQGDVHKYKIRTLNDNGGFYISPRSTFTLQELVDHYKKGNDGLCQKL 216  
Qy 188 KEPCV 192  
Db 217 SVPCM 221

RESULT 15

US-09-977-261-17  
; Sequence 17, Application US/09977261  
; Publication No. US20030054527A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: SURES, IRVINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1259  
; CURRENT APPLICATION NUMBER: US/09/977,261  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 505  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-261-17

Query Match 26.6%; Score 360.5; DB 10; Length 505;  
Best Local Similarity 41.6%; Pred No. 9.6e-27;  
Matches 77; Conservative 31; Mismatches 70; Indels 7; Gaps 2;  
Qy 12 PSPSLSSVQGGPVMTWEAERSKATAVALGSPAGGPAELSLRLGELPTIVSEDDGWTV 71  
Db 40 PGPNSHNS---NTPGIREAGSDDIIVVALDYEAHHEDLSFQKGDQWVVEESGEWKA 96  
Qy 72 LSEVSGREYNTIPSVHGVK---SHGWLYEGLSRKAEELLLPGNPGGAFLIRESOTRR 127  
Db 97 RSLATKEGYIPSNYVARVDSLETEEFKGISRKAERQLLAPGNMLGSMFMRDSEYTK 156  
Qy 128 GSYSLVSLSRPASWDRIRHYRHHCLDNGWLYISPLTFPSLQALVDHYSELADDCILL 187  
Db 157 GSYSLVSRVDYDPRQGDVHKYKIRTLNDNGGFYISPRSTFTLQELVDHYKKGNDGLCQKL 216  
Qy 188 KEPCV 192  
Db 217 SVPCM 221

Search completed: February 20, 2004, 09:50:16  
Job time : 594 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 09:18:27 ; Search time 83 Seconds  
(without alignments)  
992.172 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MSLPERRKLRPSLSVV.....RESLSFYISLNDVAVSLDDA 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1028	76.0	259	11	Q7TMX7 mus musculu
2	583	43.1	266	13	Q7T0U4
3	467.5	34.6	282	13	Q7T3X8
4	375	27.7	491	11	Q8C3I0
5	366.5	27.1	502	13	Q9DDK6
6	361.5	26.7	488	13	Q13064
7	352.5	26.1	505	4	Q961N1
8	344.5	25.5	499	11	Q8K2M8
9	340	25.1	509	6	Q95M32
10	338.5	25.0	509	11	Q91X65
11	328	24.2	509	6	Q95KR7
12	320.5	23.7	519	13	Q7ZYH5
13	315.5	23.3	537	11	Q62844
14	313.5	23.2	496	13	Q93411
15	309.5	22.9	541	11	Q99PW1
16	309	22.8	812	15	Q85466 avian sarco

17	307.5	22.7	537	13	Q7ZYK3
18	307	22.7	517	11	Q63206
19	306.5	22.7	541	11	Q8CBP1
20	306	22.6	534	4	Q16248
21	306	22.6	534	6	Q95M31
22	306	22.6	534	11	Q8K2A3
23	305.5	22.6	525	13	Q8AWF1
24	305	22.5	527	5	Q9Y121
25	304.5	22.5	517	5	Q77050
26	301	22.2	517	11	Q8BGM0
27	298.5	22.1	541	11	Q8C762
28	288.5	21.3	587	15	Q64817
29	285	21.1	511	5	Q8WQM5
30	284.5	21.0	503	5	Q8WSU5
31	283	20.9	502	13	Q8QGJ9
32	280.5	20.7	523	15	Q85477
33	280.5	20.7	526	15	Q64994
34	277.5	20.5	482	4	Q8NSD7
35	277.5	20.5	526	15	Q93080
36	276.5	20.4	535	11	Q80XU2
37	276.5	20.4	570	4	Q86VB9
38	276	20.4	535	15	Q92957
39	275.5	20.4	526	15	Q64993
40	275	20.3	546	15	Q86363
41	273.5	20.2	526	15	Q92806
42	273.5	20.2	542	11	Q93J10
43	273	20.2	504	5	Q8WSU2
44	273	20.2	545	15	Q86362
45	272	20.1	537	13	Q7ZX73

#### ALIGNMENTS

#### RESULT 1

Q7TMX7	PRELIMINARY;	PRT;	259 AA.
AC	Q7TMX7		
DT	01-OCT-2003 (TREMBLrel. 25, Created)		
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	Hypothetical protein.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6NCr; TISSUE=Hematopoietic Stem Cell;		
RX	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatane P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,		
RA	Jones S.J., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6NCr; TISSUE=Hematopoietic Stem Cell;		

```

RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC052655; AAH52655.1; -.
KW Hypothetical protein.
SQ SEQUENCE 259 AA; 28495 MW; F078DLFBB720B7CE CRC64;

Query Match 76.0%; Score 1028; DB 11; Length 259;
Best Local Similarity 79.4%; Pred. No. 1.8e-84;
Matches 208; Conservative 16; Mismatches 34; Indels 4; Gaps 3;

QY 1 MGSIPRRKSLSPSSSVQCGPVTMEASRSKATATAVALGSPFAGGPAELSLRLGEPLT 60
DB 1 MGSLSRGHT-SPPSSSGPQEPVMPQERRKVTAVALGSPFAGGPAELSLRLGEPLT 59
QY 61 IVSEDDGDMTVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKABELLLLPNGPGAF 120
DB 60 IISDDGDMTVQSEVSGREYHPSVYVAVKVAHGWLYEGLSREKABELLLLPNGPGAF 119
QY 121 RESQTRGYSYLSVLRSPASWDRIHRYIHCLDNGWLVIISPLTFPSLQALVDHYSEL 180
DB 120 RESQTRGYSYLSVLRSPASWDRIHRYIQLDNGWLVIISPLTFPSLHALVHYSEL 179
QY 181 DTICCLLKPCVLQAGPLPGKDIPLPVTVTQRTPLNWKELDSLLFSEA-ATGEESLISE 239
DB 180 DGICCPRLPCVLQKLGPLEGKDTPTPTVTPTSSLNKKLDRSLFLFAPASGEASLISE 239
QY 240 GLRESLSFYISLNDEAVSLDDA 261
DB 240 GLRESLSFYISLAED--PLDDA 259

RESULT 2
Q7T0U4 ID Q7T0U4 PRELIMINARY; PRT; 266 AA.
AC Q7T0U4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Ioshizuka S., Carninci P., Prange C.,
RA Raha S.S., Locuallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

```

---

```

RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Klein S., Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC056035; AAH56035.1; -.
KW Hypothetical protein.
SQ SEQUENCE 266 AA; 30303 MW; EF20A453A4CD673B CRC64;

Query Match 43.1%; Score 583; DB 13; Length 266;
Best Local Similarity 47.5%; Pred. No. 2.5e-44;
Matches 125; Conservative 48; Mismatches 80; Indels 10; Gaps 5;

QY 1 MGSIPRRKSLSPSSSVQCGPVTMEASRSKATATAVALGSPFAGGPAELSLRLGEPL 59
DB 1 MGNIPCKGYSIAEASLTNLEINQ-----AAADKICIFVALYNFPLSGQTDLSHFGEQL 54
QY 60 TIIVSEDDGDMTVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKABELLLLPNGPGAF 119
DB 55 NIIVSEDDGDMTVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKABELLLLPNGPGAF 114
QY 120 IRESQTRGYSYLSVLRSPASWDRIHRYIHCLDNGWLVIISPLTFPSLQALVDHYSEL 179
DB 115 IRESQTRGYSYLSVLRSPASWDRIHRYIHCLDNGWLVIISPLTFPSLQALVDHYSEL 174
QY 180 ADDICCLLKPCVLQAGPLPGKDIPLPVTVTQRTPLNWKELDSLLFSEA-ATGEESLISE 237
DB 175 ADGICCTLTPEFCVQVRV-PNPVTQRPSEPIVVRKPTLNLTWKLDSMDLFPNKDDKLNEDCPL 233
QY 238 SEGRLSELSFYISLNDEAVSLDD 260
DB 234 SLGLREAVSSYMLMTQDSSEMT 256

RESULT 3
Q7T3X8 ID Q7T3X8 PRELIMINARY; PRT; 282 AA.
AC Q7T3X8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Src-like adaptor protein.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Lebigot I., Gardellin P., Lefebvre L., Beug H., Ghysdael J.,
RA Tran Quang C.;
RT "Cloning of the avian Src-like adaptor protein cDNA."
RL Submitted (APR-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY278230; AAP34689.1; -.
SQ SEQUENCE 282 AA; 31988 MW; 85EA8B1D8986D416 CRC64;

Query Match 34.6%; Score 467.5; DB 13; Length 282;
Best Local Similarity 41.7%; Pred. No. 7.2e-34;
Matches 105; Conservative 38; Mismatches 76; Indels 33; Gaps 5;

QY 27 TWEASRSKA-----TAVA-----LCSFPAGGPAELSLRLGEPLTIVSDDGDMW 69
DB 4 TWKTFRASREGGEETINVASSTQESDFLAVLYDPSADISQPIFHVGEKLRVLSDEGGW 63
QY 70 TVLSEVSGREYNIPSVHVGKSHGWLVEGLSREKABELLLLPNGPGAFILRESQTRGS 129
DB 64 RVHSLTGTGRENIPGKYVAKVYHGWLFEGLSREKABELLQPNTKVGSFMRESSTRGL 123
QY 130 YSLVLSRSPASWDRIHRYIHCLDNGWLVIISPLTFPSLQALVDHYSELADDCCLLKE 189

```

Db 124 YLSVR-----HREYHYIFLPNNWYISPRQFQCLDLNHYSEVADGLCCVLT 177  
Qy 190 PCVLRAGPLPGKIDPLPVTQRTPLNWKEL-----DSSLFSEAATGEESLLSEGLR 242  
Db 178 PCLTQCTNNHINQVFPVVRNKNFNWRSIHRLEWGTDE---STLAAIDDSCLSVGLR 234  
Qy 243 ELSLSFISLND 254  
Db 235 ESIAYSLSLTGD 246  
RESULT 4  
Q8CE10 PRELIMINARY; PRT; 491 AA.  
AC Q8CE10;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Yamaguchi sarcoma viral.  
GN LYN  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Pancreas;  
PX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK028112; BAC25753.1; -.  
DR MGD; MGI:96892; LYN.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein-serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Ser\_Thr\_kinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR ProDom; PD000093; SH2; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00001; SH2; 1.  
DR PROSITE; PS00002; SH3; 1.  
DR PROSITE; PS00002; SH3; 1.  
SQ SEQUENCE 491 AA; 56309 MW; 2C86DB46AA4B6AAB CRC64;  
Query Match 27.7%; Score 375; DB 11; Length 491;  
Best Local Similarity 39.4%; Pred. No. 3.4e-25;  
Matches 85; Conservative 35; Mismatches 78; Indels 18; Gaps 4;  
Qy 1 MGLSPRRK-SLPSPSLSSVQ-----GGPVTWEAERSKATAVALGSPAGGPAE 50

Db 1 MGC1KSKDNLDNDEVDSTQTPVPEFLLPGQRTKDPPEQGDIVVALPYDGIHPDD 60  
Qy 51 LSLRLGELPLTVSDGDWTVLSEVSGREYNIPSVHYCKV-----SHGWLVSGLSREKAE 106  
Db 61 LSFYKGEKKVLEHGEHWKAKSLSSKEGFIPIYAKVNTLFEWFFKDIRKDSER 120  
Qy 107 LILLPGNPGGAFLIRESQTRGSYSLSVRLSRPWSWRIRHYRHLCDLNGWLYISPLTF 166  
Db 121 QLLAFGNSAGAFILRESSETLKGFSLSVRDYPHGDVHKYKIRSLDNGGYISPRITF 180  
Qy 167 PSLQALVDHYSELADDICLLKEPCVLRAGPLCK 202  
Db 181 PCISDMIKHYQKSDGLCRLEKACI-----SPKPOK 212  
RESULT 5  
Q9DDK6 PRELIMINARY; PRT; 502 AA.  
AC Q9DDK6;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Src-family tyrosine kinase SCK.  
OS Salmo salar (Atlantic salmon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.  
OX NCBI\_TaxID=8030;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hordvik I., Male R.;  
RT "A leukocyte cDNA of Atlantic salmon encoding a Src-family tyrosine  
RL kinase";  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
DR EMBL; AF321110; AAG38611.1; -.  
DR HSP; P08631; 1AD5.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00017; SH2; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR ProDom; PD000093; SH2; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00001; SH2; 1.  
DR PROSITE; PS00002; SH3; 1.  
DR PROSITE; PS00002; SH3; 1.  
KW ATP-binding; Kinase; SH3 domain; Transferase; Tyrosine-protein kinase.  
SQ SEQUENCE 502 AA; 56600 MW; 82DF0D677AA99980 CRC54;  
Query Match 27.1%; Score 366.5; DB 13; Length 502;  
Best Local Similarity 41.9%; Pred. No. 2.1e-24;  
Matches 80; Conservative 29; Mismatches 67; Indels 15; Gaps 3;  
Qy 15 SLSSVQGGQGPVTWEAERSKATAVALGSPAGGPAELSLRLGELPLTVSDGDWTVLSE 74

```

Db 45 STGPSVSDGP-----ESIALVDYEGINEGLGFKGDKLKIQESGEMWRAQSI 96
Qy 75 VSGREYNIPSVHVKVS---HGWLVEGLSREKAEELLLPGNPGGAFLIRESOTRRGSYS 131
Db 97 STGGEIPSNVYDAIDSLTEEFKVSRKDAERQLASGNKMGSMIRDSSTTKGSYS 156
Qy 132 LSVTLSPASWDRIHRHRIHCLDNGWLYISPRITFPSLQALVDHYSELADDDICLLKEPC 191
Db 157 LSVRDSQSGDVTVKHYKIRTLDNGGFYISPRITFTLQELVSHYKGLDGLCOALTSPC 216
Qy 192 VLORAGPLPGK 202
Db 217 L-----SPKPK 223

RESULT 6
Q96IN1 PRELIMINARY; PRT; 488 AA.
AC O13064
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE LYN protein tyrosine kinase.
GN LYN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Fukami Y., Funabiki K., Sato K.;
RT "Nucleotide sequence of Xenopus LYN protein tyrosine kinase.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBAJ databases.
CC 1. SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AB003358; BAA20078.1; -.
DR HSSP; P08631; IAD5.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:intracellular signaling cascade; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000093; SH2_1.
DR ProDom; PD000066; SH3_1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRKG; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00002; SH3; 1.
DR PROSITE; PS00002; SH3; 1.
DR ATP-binding; Kinase; SH3 domain; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 488 AA; 5794 MW; E7E70668B6EA92B2 CRC64;

Query Match 26.7%; Score 361.5; DB 13; Length 488;
Best Local Similarity 38.3%; Pred. No. 5.6e-24;
Matches 82; Conservative 79; Mismatches 79; Indels 17; Gaps 4;

Qy 1 MCSLSRR-----KSLPSLSSSVQGGPVTMEARSKATAVALGSPAGPAELS 52

```

```

Db 1 MGC1KSTNDPRDANGKKNPASRTPSLLPGQ-KW1QDIEQGNIIVIALPYQGIHEDDLS 59
Qy 53 LRLGEPLTIVSDGDMWTVLSEVSGREYNIPSVHVKV-----SHGWLVEGLSREKABELL 108
Db 60 FKKGKELKLEHGEHWNKAKSLSTKKEGFIPTSNVARVNTLETEENFFKOLTRKDAERQL 119
Qy 109 LLPGNPGAFITRESOTRRGSYSLSVLSRSPASWDRIHRHRIHCLDNGWLYISPRITFPS 168
Db 120 LAPGNPGAFILRESOTRRGSYSLSIRDCDPTQGDVHKYKIRTLDNGGYISPRITFTS 179
Qy 169 LQALVDHYSELADDDICLLKEPCVLRAGPLPGK 202
Db 180 INEMIQHYQKQADGLCRKLDKFCF-----SPKPK 209

RESULT 7
Q96IN1 PRELIMINARY; PRT; 505 AA.
AC Q96IN1
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein (B lymphoid tyrosine kinase).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RT TISSUE=Blood, and Lymph;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBAJ databases.
DR EMBL; BC007371; AAH07371.1; -.
DR EMBL; BC032413; AAH32413.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:intracellular signaling cascade; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000093; SH2_1.
DR ProDom; PD000066; SH3_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00002; SH3; 1.
DR PROSITE; PS00002; SH3; 1.
DR Hypothetical protein; ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 505 AA; 57706 MW; B5F739BEF8389176 CRC64;

Query Match 26.1%; Score 352.5; DB 4; Length 505;
Best Local Similarity 43.6%; Pred. No. 3.8e-23;
Matches 75; Conservative 24; Mismatches 68; Indels 5; Gaps 2;

Qy 25 PVTMEARSKATAVALGSPAGPAELSRLGLCEPLTIVSDGDMWTVLSEVSGREYNIPS 84
Db 51 PPEHLEDDKGFVALYDYTMNDRDLQMLKGEKQLVKGTDGWLARSLLVTGREGYVPS 110
Qy 85 VHYGVKVS---HGWLVEGLSREKAEELLLPGNPGGAFLIRESOTRRGSYSLSVLSRPA 140
Db 111 NFVARVESLEMEERWFFRSQGRKEARQLLAPINKAGSFLIRESETNKGAFSVK-DVTT 169
Qy 141 SWDIRHRIHCLDNGWLYISPRITFPSLQALVDHYSELADDDICLLKEPCV 192

```

Query March	25.14; Score 140; DB 6; Length 509;
Best Local Similarity	40.64; Pred. No. 5.1e-22;
Matches 73; Conservative	26; Mismatches 71; Indels 10; Gaps 2
QY	26 VTMEAKRSKAT-----AVALGSPFAGGPAEELSRLGELPTIVSDGDMWTVLSEVSGRE 79
Db	49 VTVEGSPNPASPLQDNLVIALHSEYPSHDGDLGFKEKGELRLLEQSGEWMKAQSLTTGQE 108
QY	80 YNIPSFVHGKVS----HGWLVEGSLREKABELLLLPNGCGAFLIREQCTRGRGYSLSVR 135

Db 109 GFIPNFVAKANSLEPEPFFKNLSKDAERQQLAPGNTGSHFLRESSTAGSFSLSVR 168  
 Q91X65  
 Q91X65 PRELIMINARY; PRT; 509 AA.  
 AC 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to lymphocyte-specific protein tyrosine kinase (2 days neonate  
 thymus thymic cells cDNA, RIKEN full-length enriched library,  
 DE clone:E430002B06 product:lymphocyte protein tyrosine kinase, full  
 DE insert sequence).  
 GN LCK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RA Strausberg B.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katon H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Nishi K., Nomura K., Numazaki R., Okazaki K., Sakai K., Sakazume N., Sano H.,  
 RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akai H., Tanaka Y., Tanaka T.,  
 RA Tomaru A., Taya T., Yasunishi A., Yamatsuta M., Hayashizaki Y.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RA The FANTOM Consortium;  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs";  
 RL Nature 420:563-573 (2002).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690 (2001).  
 [5]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning";  
 RL Meth. Enzymol. 303:19-44 (1999).  
 [6]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes";  
 RL Genome Res. 10:1617-1630 (2000).

RN SEQUENCE FROM N.A.  
 RP STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagata S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Tanaka T., Matsura S., Kawai J.,  
 RA Yoneda Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer";  
 RL Genome Res. 10:1757-1771 (2000).  
 DR EMBL; BC011474; AAH11474.1;  
 DR EMBL; AK088001; BAC40086.1;  
 DR F1R; I48845; I48845;  
 DR MGD; MGI:96756; Lck.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH2DOMAIN.  
 DR PRINTS; PR00103; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00252; SH3; 1.  
 DR SMART; SM00219; TyrcK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00001; SH2; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 509 AA; 57942 MW; 3513102F49A7FD0B CRC64;  
 Query Match 25.0%; Score 338.5; DB 11; Length 509;  
 Best Local Similarity 39.1%; Pred. No. 7e-22;  
 Matches 77; Conservative 28; Mismatches 85; Indels 7; Gaps 2;  
 QY 3 SLPSRKSLPSLSVSGREYNIPSVHYGKVS----HGWLYEGLSREKAEELLILPGNPGGAP 118  
 DB 35 SLPIRNGSEVRDPL---VYEGSLPPASPLQDNLVIALHSYEPHSDGDLGPEKGEQLIL 91  
 QY 63 SDGQGWTVLSEVSGREYNIPSVHYGKVS----HGWLYEGLSREKAEELLILPGNPGGAP 118  
 DB 92 EQSGEWWKAQSLITGEGFIPENFVAKANSLEPEPFFKNLSKDAERQQLAPGNTGSHF 151  
 QY 119 LRESQTRGSGVSLVRLSRPASWDRIRHYRHCIDNGWLYISPELTSPSLQALVDHYSE 178  
 DB 152 LIRESSESTAGSFSLSVRDFDQNGGEVWVHYKIRNLDNGGFIYSPIRITPFLGLHDLVRYHTN 211  
 QY 179 LADDICLLKKEPCVLQR 195  
 DB 212 ASDGLCTKLSRPCQIQK 228  
 RESULT 11  
 Q95KR7 PRELIMINARY; PRT; 509 AA.  
 ID Q95KR7  
 AC Q95KR7;

DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Lck tyrosine kinase (Fragment).  
 GN LCK.  
 OS Saimiri sciureus (Common squirrel monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.  
 OX NCBI\_TaxID=9521;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21424508; PubMed=11531187;  
 RA Greve T, Tamguney G, Fleischer B, Fickenscher H, Brooker B.M.;  
 RT "Downregulation of p56lck tyrosine kinase activity in T cells of  
 squirrel monkeys (Saimiri sciureus) correlates with the non-  
 transforming and apathogenic properties of herpesvirus saimiri in its  
 natural host";  
 RT J. Virol. 75:9252-9261(2001).  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL: AJ277921; CAC38871.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR000980; SH2.  
 DR InterPro: IPR001452; SH3.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR InterPro: IPR008266; Tyr\_pkinase\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR PRINTS: PRO0401; SH2DOMAIN.  
 DR PRINTS: PRO0452; SH3DOMAIN.  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR ProDom: PD000093; SH2; 1.  
 DR ProDom: PD000066; SH3; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR SMART: SM00326; SH3; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS0011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN KINASE\_TYR; 1.  
 DR PROSITE: PS50001; SH2; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 KW ATP-binding; Kinase; SH3 domain; Transferase; Tyrosine-protein kinase.  
 FT NON\_TER 509 509  
 SQ SEQUENCE 509 AA; 58253 MW; 5088C6407D109519 CRC64;  
 Query Match 24.2%; Score 328; DB 6; Length 509;  
 Best Local Similarity 41.4%; Pred. No. 6.2e-21;  
 Matches 67; Conservative 27; Mismatches 64; Indels 4; Gaps 1;  
 QY 38 VALGSPFAGGPAELSLRLGEPLTIYSEDGDWTVLSEVSGREYNIPSVHVGKVS-----HG 93  
 Db 67 IALHSTPEHDGDLGPEKEGHELRILEQNGEWKASLTIGQGFFVFFVAKANLEPEP 126  
 QY 94 WLYEGLSREKABELLALPGNPGGAFLIRESQTRGSYSLSVRLSRPASWDRIHYRIHCL 153  
 Db 127 WFFKNLSRKAERQLLAPGNTGHSFLIRESESTAGSFSLSVRDPDQNGEVVKYKIRNL 186  
 QY 154 DNGWLYISRLTFPSLQALVDHYSLEADICCLKEPCVLOR 195  
 Db 187 DNGGFVISPRITFSGLHVLVRHYTNASDGLCTRLSRPQCTQK 228  
 RESULT 12  
 Q7ZYH5  
 ID Q7ZYH5 PRELIMINARY; PRT; 519 AA.  
 AC Q7ZYH5  
 DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Similar to Fyn proto-oncogene.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC043783; AB43783.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004713; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR001452; SH2.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR InterPro: IPR008266; Tyr\_pkinase\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR PRINTS: PRO0401; SH2DOMAIN.  
 DR PRINTS: PRO0452; SH3DOMAIN.  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR ProDom: PD000093; SH2; 1.  
 DR ProDom: PD000066; SH3; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR SMART: SM00326; SH3; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS50011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN KINASE\_TYR; 1.  
 DR PROSITE: PS50001; SH2; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 SQ SEQUENCE 519 AA; 58486 MW; A790E29424AE7037 CRC64;  
 Query Match 23.7%; Score 320.5; DB 13; Length 519;  
 Best Local Similarity 36.2%; Pred. No. 3e-20;  
 Matches 75; Conservative 37; Mismatches 76; Indels 19; Gaps 5;  
 QY 4 LPSRRKSLPSFLSSVQG--QGP-----VTMEARSKATAVALGSPAGGPAELSL 53  
 Db 33 MPFTQETHPTFTIGIGVMSGSPHRGMGVGTAGA---TLFIALYDYDATEDDLTF 88  
 QY 54 RLGEPTIV-SEGDWMTVLSEVSGREYNIPSVHVGKVS-----SHGWLYEGLSREKABELL 108  
 Db 89 QKGEKPHIINSEGDWEARSLSSGNTGYPSNYAPMNSIQSKDWFFGKIGKDAERLL 148  
 QY 109 LLPNPGGAFLIRESQTRGSYSLSVRLSRPASWDRIHYRIHCLDNGWLYISRLTFPS 168  
 Db 149 LSEGNPGTGMIRESETTKGAYLSLVRDWDASGDHAKYKIRKLSGGFYITTRVQFNT 208  
 QY 169 LQALVDHYSLEADICCLKEPCVLOR 195  
 Db 209 VEELVEHYGFNDGLCHLLTSPCINQK 235  
 RESULT 13  
 Q62844  
 ID Q62844 PRELIMINARY; PRT; 537 AA.  
 AC Q62844  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)



Search completed: February 20, 2004, 09:32:11  
Job time : 90 secs

Query Match	22.98;	Score	309.5;	DB	11;	Length	541;
Best Local Similarity	32.08;	Prod. NO.	3.2e-19;				
Matches	81;	Conservative	#41;	Mismatches	92;	Indels	39;
						Gaps	7;
38	QY	VALGSFAGGPAELSLRLGELPLTVISE-DGDWMTVLSEVSREYNIPSVHVGVK-----SH	92				
95	Db	VALYDEYARTTDLSPKKGERRFQIINTEGDNWEARSATGNGVYIPSNYVAPADSIQAE	154				
93	QY	CWLVEGSLREKABELLLLPGNPGCALFIBESQTRGSGYSLSVRLSRPASWDEIR-----H	147				
155	Db	EWYFGKGRKDAERLLNPGNQGFIFLVRESEITTKGYSLSIR-----DWDEVGDNVNH	209				
148	QY	YRIECLONGWLYISPRITFESLQALVDHYSELADICCLKEPC-----VLQAGP	198				
210	Db	YKIRKLDNGGYIITRAQFPDTLQKLVKGVTEHADGLCHKLTTVCPTVKPQTQGLKDAWE	269				
199	QY	LPGKXDILPVTVQR-----FPLNWKELDSSLFSEATGESILSEGLRES	244				
270	Db	IPRSRLREVKLGGCGEYVWGTNGTTKVAIKTLKPTWMPFELQEAQIMKLRHDK	329				
245	QY	L-SFYISLNDEAV	256				
330	Db	LVPLVAVVSEBPI	342				

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:34:08 ; Search time 3092 Seconds  
(without alignments)  
3658.645 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSILPSRKSLPSPSLSSV.....RESLSFVISLNDRAVSLDDA 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-C=/cgn2\_1/uspto.spool/p/US09939853/runat\_19022004\_145339\_24464/app.query.fasta\_1.455  
-DB=GenEmbl -QWTF=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LQOPCL=0 -LQOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09939853\_@CGN\_1\_1\_5265\_@runat\_19022004\_145339\_24464 -NCPU=6 -ICPU=3  
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sv:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C	1	1353	100.0	1183	6	AX443133 Sequence
	2	1353	100.0	1183	6	AX443135 Sequence
	3	1347	99.6	786	6	AX511153 Sequence
	4	1347	99.6	786	6	AX572845 Sequence
	5	1347	99.6	786	9	AF290985 Homo sapi
	6	1347	99.6	786	9	AF326353 Homo sapi
	7	1347	99.6	2415	9	AK025645 Homo sapi
	8	1347	99.6	2538	9	BC042041 Homo sapi
	9	1347	99.6	2567	6	AX452880 Sequence
	10	1341	99.1	2788	6	AX780857 Sequence
C	11	1196.5	88.4	737	6	AX511155 Sequence
	12	1196.5	88.4	737	9	AF290986 Homo sapi
	13	1028	76.0	777	6	AX511151 Sequence
	14	1028	76.0	1348	6	AX511150 Sequence
	15	1028	76.0	1348	10	AF287467 Mus muscu
	16	1028	76.0	1384	10	BC052655 Mus muscu
	17	1019	75.3	1321	10	AF434990 Mus muscu
	18	583	43.1	1571	5	BC056035 Xenopus l
	19	501	37.0	1631	10	AY217759 Rattus norv
	20	492	36.4	921	10	MMU29056 Mus musculu
	21	492	36.4	2648	10	BC032922 Mus muscu
	22	492	36.4	2773	10	MMU131777 Mus muscu
	23	491.5	36.3	841	10	AY079449 Mus muscu
	24	491.5	36.3	841	10	AY079450 Mus muscu
	25	487	36.0	1076	9	HSU30473 Homo sapien
	26	487	36.0	1870	9	BC007042 Homo sapi
	27	487	36.0	2021	9	HSU44403 Human Src-1
	28	487	36.0	2109	6	AX428893 Sequence
	29	487	36.0	2665	6	AX333017 Sequence
	30	487	36.0	2665	9	D89077 Homo sapien
	31	487	36.0	3090	9	HSM808530 Homo sapi
	32	482	35.6	145068	9	HSU4977B1 Human DNA
	33	482	35.6	145833	2	AC026539 Homo sapi
	34	467.5	34.6	849	5	AY278230 Gallus ga
	35	380.5	28.1	2703	10	MUSLYNB Rattus norv
36	375.5	27.8	1538	10	RATLYNBYR Rattus norv	
37	375.5	27.8	2302	10	RNAF000301 Rattus norv	
38	375.5	27.8	2824	10	RNAF000302 Rattus norv	
39	370.5	27.4	1539	12	AY335782 Synthetic	
40	370.5	27.4	2298	6	AR380613 Sequence	
41	370.5	27.4	2298	6	AX779985 Sequence	
42	370.5	27.4	2298	9	HUMLYN Human lyn m	
43	370.5	27.4	4175	6	AX779984 Sequence	
44	370	27.3	1601	10	RATLYNBYR Rattus norv	
45	370	27.3	2041	10	RNAF000300 Rattus norv	

ALIGNMENTS

RESULT 1

AX443133	AX443133	1183 bp	DNA	linear	PAT 02-JULY-2002
LOCUS	Sequence 74 from Patent WO0216599.				
DEFINITION	AX443133				
ACCESSION	AX443133				
VERSION	AX443133.1	GI:21690555			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R., Shinkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E., Topper, J.N. and Yang, R.B.				
TITLE	Proteins and nucleic acids encoding same				
JOURNAL	Patent: WO 0216599-A 74 28-FEB-2002;				
FEATURES	Curagen Corporation (US); COR THERAPEUTICS, INC. (US)				
source	Location/Qualifiers				
	1..1183				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Alignment Scores:					
Pred. No.:	2,01e-102	Length:	1183		
Score:	1353.00	Matches:	261		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-939-853A-75 (1-261) x AX443133 (1-1183)					
Qy	1 MetGlySerLeuProSerArgArgIySserLeuProSerProSerLeuSerSerVal	20			
Db	398 ATGGGAAGTCTCCACAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTCTGTC	457			
Qy	21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40			
Db	458 CRAAGGCCAGGACCTGTGACCATGGAGCAGAGAGAGCAAGCCACGCCGTGGCCCTG	517			
Qy	41 GlySerPheProAlaGlyClyProAlaGluLeuSerLeuArgIleuGlyGluProLeuThr	60			
Db	518 GGCAAGTTTCCCGGACAGGTGGCCGCGCAGCTGCTGCTGAGATCTGGGAGGCATGACC	577			
Qy	61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr	80			
Db	578 ATCGTCTCTGAGGATCGAGACTGGTGGACGGTGTCTGAACTCTCAGCAGAGAGTAT	637			
Qy	81 AsnIleProSerValHisValGlyIySValSerHisGlyTrpLeuTyrGluGlyLeuSer	100			
Db	638 AAGATCCCGAGCGTCACGTGGGCAAGCTCCCATGGTGGCTGTATGAGGCGCTGAGC	597			
Qy	101 ArgGlyIySAlaGluLeuLeuLeuLeuProGlyAsnProGlyAlaPheLeuIle	120			
Db	698 AGGAGAAAGCAGAGAACTGCTGTGTACCTGGGAACCTCGAGGGGCTTCTCTCATC	757			
Qy	121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgIleuSerArgProAla	140			
Db	758 CCGGAGAGCCAGACACAGAGAGGCTTCTTACTCTCTGTGTCGCTGAGCGCCCTGCA	817			
Qy	141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160			
Db	818 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGCAATGGCTGGCTGTACATC	877			
Qy	161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180			
Db	878 TCACCGGCGCTCACCTTCCCTTCACTCAGAGCCCTGGTGGACCATCTACTGTAGCTGGCG	937			
Qy	181 AspAspIleCysCysLeuLeuGlyGluProCysValLeuGlnArgAlaGlyProLeuPro	200			
Db	938 GATGACATCTGCTGCTTCTCAAGAGAGCCTGTGCTCTGACAGGCTGGCCGCTCCCT	997			

Db 486 AGGAGAAAGCAGAGGAACCTCTGTGTTACCTGGGAACCCCTGGAGGGCTTCTCTCATC 427  
Qy 121 ArgGluSerGlnThrArgArgGlySerTyfSerLeuSerValArgLeuSerArgProAla 140  
Db 426 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTCAAGTCCGCCCTCAGCCGCCCTGCA 367  
Qy 141 SerTrpAspArgIleArgHisTyfArgIleHisCysLeuAspAsnGlyTrpLeuTyfIle 160  
Db 366 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACATGCTGGCTGTATCATC 307  
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyfSerGluLeuAla 180  
Db 306 TCACCGGCCCTCACTTCCCTCTCACTCAGGCCCTGGTGGACCACTTACTCTGAGCTGGCG 247  
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
Db 246 GATGACATCTGCTGCTACTCAAGAGGCCCTGTGCTCGAGAGGCTGGCCGCCCTCCCT 187  
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
Db 186 GGCAGAGATATACCCCTACCTGTGACTGTGACAGAGGACACCACTCAACTGGAAAGAGCTG 127  
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
Db 126 CACAGCTCCCTCTGTTTCTGAGCTGCCACAGGGAGGAGTCTCTCTCAGTGAGGGT 67  
Qy 241 LeuArgGluSerLeuSerPheTyfIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
Db 66 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGAT 7  
Qy 261 Ala 261  
Db 6 GCC 4  
RESULT 3  
AX511153  
LOCUS AX511153 786 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 4 from Patent WO0242452.  
ACCESSION AX511153  
VERSION AX511153.1 GI:23392046  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS McGlade, J.C. and Loreto, M.P.  
TITLE Adapter gene  
JOURNAL Patent: WO 0242452-A 4 30-MAY-2002;  
The Hospital for Sick Children (CA)  
FEATURES  
Location/Qualifiers  
Source 1. 786  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3 9e-102 Length: 786  
Score: 2347.00 Matches: 260  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.56% Indels: 0  
DB: 6 Gaps: 0  
US-09-939-853A-75 (1-261) x AX511153 (1-786)  
Qy 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
Db 1 ATGGGAAGTCTCCAGAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC 60  
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40

Db 61 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGCAAGGCCACAGCCGTGCCCTG 120  
Qy 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGluProLeuThr 60  
Db 121 GGCAGTTCCTCCGGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGGCATGACC 180  
Qy 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyf 80  
Db 181 ATCGTCTCTGAGATGGAGACTGGTGGACGGTGTCTGAAATCTCAGGCAGAGAGTAT 240  
Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyfGluGlyLeuSer 100  
Db 241 AACATCCCGCAGGTCCAGTGGCCCAAGTCTCCCATGGGTGGCTGTATGAGGCCCTGAGC 300  
Qy 101 ArgGluLysAlaGluGlnLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120  
Db 301 AGGAGAGAAGCAGAGAACTGCTGTGTGTACTCGGAACCCCTGGAGGGCTTCTCTCATC 360  
Qy 121 ArgGluSerGlnThrArgArgGlySerTyfSerLeuSerValArgLeuSerArgProAla 140  
Db 361 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTGAGTCCGCCCTCAGCCGCCCTGCA 420  
Qy 141 SerTrpAspArgIleArgHisTyfArgIleHisCysLeuAspAsnGlyTrpLeuTyfIle 160  
Db 421 TCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACAATGGCTGGTGTATCATC 480  
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyfSerGluLeuAla 180  
Db 481 TCACCGCGCTCCTCCTTCCCTCCTCAGTCCAGGCCCTGGTGGACCATTAATCTCAGCTGGCG 540  
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
Db 541 GATGACATCTGCTGCTACTCAAGAGGCCCTGTGCTGACAGAGGCTGGCCGCCCTCCCT 600  
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
Db 601 GGCAGAGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGGAAAGAGCTG 660  
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
Db 661 GACAGCTCCCTCTGTTTCTGAGCTGCCACAGGGAGGAGTCTCTCTCAGTGAGGGT 720  
Qy 241 LeuArgGluSerLeuSerPheTyfIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGGCTGTCTCTTTGGATGAT 780  
Qy 261 Ala 261  
Db 781 GCC 783  
RESULT 4  
AX572845  
LOCUS AX572845 786 bp DNA linear PAT 29-NOV-2002  
DEFINITION Sequence 1 from Patent WO02055707.  
ACCESSION AX572845  
VERSION AX572845.1 GI:26004935  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Holland, S.J., Mendenhall, M.K., Pardo, J., Spencer, C., Fu, A.C.,  
Luo, Y., Payan, D.G., Mancebo, H.S., Wu, J., Zhou, X., Shen, M.,  
Liao, X.C. and Sheng, N.  
TITLE Cloning of an inhibitor of antigen-receptor signaling by a  
retroviral-based functional screen  
JOURNAL Patent: WO 02055707-A 1 18-JUL-2002;  
Rigel Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
Source 1. 786  
/organism="Homo sapiens"

```
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..786
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD58542.1"
/db_xref="GI:26004936"
/db_xref="RENTREMBL:CAD58542"
/translation="MGSLSRRKSLPSLSVSSVQGGPVVMEARSKATAVALGSPF
AGPAELSLRLGPELTVISDGDWMTVLSEVSGREYNIPSVHAKVSHGWLVEGLSRE
KAELLPLPGNPGAFILRESQTRRSYLSVLSRSPASMDRIHRYIHCLDNGWLYI
SPRLTFPSLQALVDHYSELADDICLLKEPCVLQRAGPLPGKDIPLPVTVORTPLNWK
ELDSLLFSEAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"
ORIGIN
Alignment Scores:
Pred. No.: 3.9e-102 Length: 786
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.56% Indels: 0
DB: 6 Gaps: 0
US-09-939-853A-75 (1-261) x AX572845 (1-786)
Qy 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGATTCCTCTGTC 60
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGCCAGGAGACCTGTGACCATGGAAGCAGAGAGAGCAAGCCAGCCGTCGGCCCTG 120
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLysGluProLeuThr 60
Db 121 GGCAGTTCCTCCGAGGTGGCCGCGAGCTGTCTGAGACTCGGGAGCCATTGACC 180
Qy 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCTCTCTCAGGATGGAGACTGTGTGAGCGTGTCTGAAGTCTCAGCAGAGAGTAT 240
Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCGAGCTCCAGTGGCCGCGCAAGTCTCCATGGTGGTGTGTATGAGGCGCTGAGC 300
Qy 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120
Db 301 AGGAGAAAGCAGAGGAACCTGCTGTGTACCTGGGAACCTGGAGGGGCTTCTCATC 360
Qy 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGCAGCAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCCCTGCA 420
Qy 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db 421 TCCTGGGACCGGATCAGACACTCAGGATCCACGCTTCAATGCTGATGCTGGCTGATCATC 480
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 481 TCACCGCGCCCTCACCTCCCTCTACTCCAGGCCCTGTGTCTCTGAGCCATTACTCTGAGCTGGC 540
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 541 GATGACATCTCTGCTACTCAAGGAGCCCTGTGTCTCTGAGAGGGCTGGCCCGCTCCCT 600
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLeu 220
Db 601 GGCAGGATATACCCCTACCTGTGACTGTGACAGAGGACACCACTCAACTGGAAGAGCTG 660
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 661 GACAGTCCCTCTCTGTTTCTGAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGT 720
Qy 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
Db 721 CTTCCGGAGTCCCTCAGCTTCTACATCAGCTGAATCAGCAGGCTGTCTTTTGGATGAT 780
Qy 261 Ala 261
Db 781 GCC 783
RESULT 5
AF290985 786 bp mRNA linear PRI 21-JAN-2003
LOCUS Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
DEFINITION Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF290985
VERSION AF290985.1 GI:17351920
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLES Cloning and characterization of human Src-like adaptor protein 2
and a novel splice isoform, SLAP-2-v
JOURNAL Oncogene 22 (2), 266-273 (2003)
MEDLINE 22415750
PUBMED 12527895
REFERENCE 2 (bases 1 to 786)
AUTHORS Loreto, M.P. and McGlade, C.J.
TITLES Direct Submission
JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
source
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1..786
/notes="SLAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAL38197.1"
/db_xref="GI:17351921"
/translation="MGSLSRRKSLPSLSVSSVQGGPVVMEARSKATAVALGSPF
AGPAELSLRLGPELTVISDGDWMTVLSEVSGREYNIPSVHAKVSHGWLVEGLSRE
KAELLPLPGNPGAFILRESQTRRSYLSVLSRSPASMDRIHRYIHCLDNGWLYI
SPRLTFPSLQALVDHYSELADDICLLKEPCVLQRAGPLPGKDIPLPVTVORTPLNWK
ELDSLLFSEAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"
CDS
ORIGIN
Alignment Scores:
Pred. No.: 3.9e-102 Length: 786
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.56% Indels: 0
DB: 9 Gaps: 0
US-09-939-853A-75 (1-261) x AF290985 (1-786)
Qy 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGATTCCTCTGTC 60
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGCCAGGAGACCTGTGACCATGGAAGCAGAGAGAGCAAGCCAGCCGTCGGCCCTG 120
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLysGluProLeuThr 60
Db 121 GGCAGTTCCTCCGAGGTGGCCGCGAGCTGTCTGAGACTCGGGAGCCATTGACC 180
```

```
QY 61 lIeValSerGluAaspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCGTCTCTGAGGATGGAGATGGTGGACGGTCTGTCTGAAGTCTCAGGACAGAGATAT 240
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCCAGCGTCCACGTCGCCAAAGTCTCCATGGGTGGTGTATGAGGGCTGAGC 300
QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLe 120
Db 301 AGGGAGAAAGCAGAGAACTGCTGTGTACTGGGAACCCCTGGAGGGCTTCTCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGACCCAGAGAGGCTTCTTACTCTGTCTGAGTCCGCTCAGCGCCCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTGTGACAATGGCTGCTCATC 480
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 481 TCACCGCGCTCACCTTCCCTCATCTCAGGCCCTGGTGGACCATTACTCTGAGCTGGCG 540
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 541 GATGACATCTGCTGCTACTCAAGAGCGCTGTGTCTCTGACAGGGCTGGCGCCCTCCCT 600
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 601 GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCACTGGAAAGAGCTG 660
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGluGly 240
Db 661 GACAGCTCCCTCTCTTCTCAAGCTCCACAGGGAGAGTCTCTTCTCAGTGAAGGCT 720
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGAAATGAGAGGCTGTCTCTTTGGATGAT 780
QY 261 Ala 261
Db 781 GCC 783
RESULT 6
AF326353 786 bp mRNA linear PRI 08-NOV-2001
LOCUS Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
DEFINITION AF326353
ACCESSION AF326353
VERSION AF326353.1 GI:16797891
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 786)
AUTHORS Holland,S.J., Liao,X.C., Mendenhall,M.K., Zhou,X., Pardo,J.,
Chen,P., Spencer,C., Fu,A.C., Sheng,N., Yu,P., Pali,E., Nagin,A.,
Shen,M., Yu,S., Chan,S., Wu,X., Li,C., Woetschlager,M., Y.,
Aversa,G., Kolbinger,F., Bennett,M.K., Mollineaux,S., Luo,Y.,
Payan,D.G., Mancebo,H.S.Y. and Wu,J.
Functional cloning of Src-like Adaptor Protein-2 (SLAP-2), a Novel
Inhibitor of Antigen Receptor Signaling
J. Exp. Med. 194 (9), 1263-1276 (2001)
TITLE 2153259
JOURNAL MEDLINE
PUBMED 1169592
REFERENCE 2 (bases 1 to 786)
AUTHORS Holland,S.J., Mendenhall,M.K., Zhou,X., Spencer,C., Pardo,J.,
Fu,A.C., Sheng,N., Shen,M., Liao,C., Luo,Y., Payan,D.G.,
Mancebo,H.S.Y. and Wu,J.
Direct Submission
TITLE Submitted (05-DEC-2000) Rigel Pharmaceutical Inc., 240 East Grand
JOURNAL Avenue, South San Francisco, CA 94080, USA
```

```
FEATURES
source
Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
CDS
1..786
/note="SLAP2; SH2 and SH3 domain-containing adaptor"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAL29204.1"
/db_xref="GI:16797892"
/translation="MGSLPSPRRKSLPSLSLSSVQGGPVTWEERSKATAVALGSGFP
AGPRLSRLGEPLTIVSDGDMTWLSEVSGREYNIPSVHVAKVSHGWLVEGJSRE
KAEILLIPGNPGGAFILRESQTRRGSYLSVRLSPASWDRIYRIHCLDNGWLYI
SRPLTFPSLQALVDHYSELADICLLKPECVLQRAGPLGKDIPLPVTQRTPLNWK
ELDSSLLFSEATGSESLSEGLRESLSFYISLNDLSEAVSLDDA"
ORIGIN
Alignment Scores:
Pred. No.: 3,9e-102 Length: 786
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.56% Indels: 0
DB: 9 Gaps: 0
US-09-939-853A-75 (1-261) x AF326353 (1-786)
QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 1 ATGGGAGTCTGCCACAGCAGAGAAATCTTGCCACAGCCCAAGCTTGAGTTCTCTGTC 60
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 61 CAAGGGCCAGGAGCTGTGACCATGGAAGCAGAGAGAGCAAGCCACACCGGTGGCCCTG 120
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 121 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
QY 61 IleValSerGluAaspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 181 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTCTGAAAGTCTTCAGGACAGAGAT 240
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100
Db 241 AACATCCCGAGCTCCACGTGGCCAAAGTCTCCATGGGTGGTGTATGAGGGCTGAGC 300
QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLe 120
Db 301 AGGGAGAAAGCAGAGAACTGCTGTGTACTGGAACCCCTGGAGGGCTTCTCTCATC 360
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 361 CGGAGAGCCAGACCCAGAGAGGCTTCTTACTCTGTCTGAGTCCGCTCAGCGCCCTGCA 420
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
Db 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTGTGACAATGGCTGCTCATC 480
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 481 TCACCGCGCTCACCTTCCCTCATCTCAGGCGCTGTGCTGAGAGGGCTGGCCCGCTCCCT 540
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 541 GATGACATCTGCTGCTACTCAAGAGCGCTGTGTCTCTGACAGGGCTGGCGCCCTCCCT 600
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
Db 601 GGCAGAGATATACCCCTACCTGTGACTGTGAGAGGACACCACTCACTGGAAAGAGCTG 660
```

```

QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGly 240
Db 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTCTTCTCAGTGAGGT 720

QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
Db 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGNAATCAGCAGGCTGCTCTTTGGATGAT 780

QY 261 Ala 261
Db 781 GCC 783

RESULT 7
AK025645 2415 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA: FLJ21992 fis, clone HEP06554.
DEFINITION AK025645
VERSION AK025645.1 GI:10438227
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Kawabata,A., Hiki,j., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2415)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source Location/Qualifiers
1..2415
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP06554"
/cell_lines="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/notes="cloning vector pME18SFL3"
61..846
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB15201.1"
/db_xref="GI:10438228"
/translation="MGSLPGRKSLPSPSSVSSVQGGPVVMEABRSKATAVALGSPFP
AGPAELSLRIGLEPLTVSEGDWMTVLSEVSGREYNIPSHVAKYSHGWLVEGLSRE
KAEELLPLPGPGAFIRIESQTRGSYLSVLRSPASWDRIHRIHICLDNGWLYI
SPRTFFPSLQALVDHYSELADIDICLLKEPCVLQRAGLPKDIPLVTVQRTPLNWK
ELDSLFLFSEATGEESLLSGLSELSFVLSLNDEAVSLDDA"

ORIGIN
Alignment Scores:
Pred. NO.: 1.44e-101 Length: 2415
Score: 1347.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.56% Indels: 0

```

```

DB: 9 Gaps: 0
US-09-939-853a-75 (1-261) x AK025645 (1-2415)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
Db 61 ATGGGAAGTCTGCCAGCAGAGAAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 120

QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluSerGlyAlaThrAlaValAlaLeu 40
Db 121 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAAGCAAGGCCACAGCCGTGGCCCTG 180

QY 41 GlySerPheProAlaGlyProAlaGluLeuSerLeuArgGlyGluProLeuThr 60
Db 181 GGCAGTTCCTCCGCGCAGGTGGCCGCGCAGCTGTGCTGAGACTCGGGAGCCATTGACC 240

QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyValGluTyr 80
Db 241 ATCGTCTCTGAGGATGGAGACTGGTGGACGTGTGTCTGAAGTCTCAGGAGAGAT 300

QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 301 AACATCCCGCAGCTCCAGCTGGCCAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGC 360

QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 361 AGGGAGAAGCAGAGGAACTGCTGTGTGTACTCTGGAAACCTCGAGGGGCTTCTCTCATC 420

QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 421 CGGGAGAGCCAGACCCAGGAGAGGCTCTTACTCTCTGTCTCAGTCCCGCTCAGCGCCCTGCA 480

QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db 481 TCCTGGGACCCGATCAGACACTACAGGATCCACTGCCTTGACATGGCTGGCTGTACATC 540

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
Db 541 TCACCGCGCTCACCTTCCCTCTCCTCAGGCCCTGGTGGACCATTACTCTGAGCTGGCG 600

QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 601 GATGACATCTGTGCTACTCAAGGAGCCCTGTGTCTGTCAGAGGGCTGGCCCGCTCCT 660

QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLeu 220
Db 661 GGCAGGAGATATACCCCTACCTGTGACTGTGAGAGGACCCACTCACTTGAAGAGCTG 720

QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGly 240
Db 721 GACAGCTCCCTCTCTGTTTCTTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGT 780

QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
Db 781 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGGAATGACGAGGCTGTCTCTTTGGATGAT 840

QY 261 Ala 261
Db 841 GCC 843

RESULT 8
BC042041 2538 bp mRNA linear PRI 07-OCT-2003
LOCUS Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (cdna
clone MGC:49845 IMAGE:4429896), complete cds.
DEFINITION BC042041
ACCESSION BC042041.1 GI:27469842
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2538)
REFERENCE
1 (bases 1 to 2538)

```

## AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toehiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bobak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Buttenfield, Y.S., Krzywinski, M.I., Skalska, J., Smalios, D.E., Schnerch, A., Schein, J.B., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)

22388257

12477932

2 (bases 1 to 2538)

Strausberg, R.

Direct Submission

Submitted (23-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-sngc.stanford.edu>

Contact: (Dickson, Mark) [mcdepaxil.stanford.edu](mailto:mcdepaxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAP Plate: 88 Row: a Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28416422.

## FEATURES

## source

Location/Qualifiers

1..2538

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:49845 IMAGE:4429896"

/tissue type="Prostate, adenocarcinoma."

/clone.lib="NIH MGC\_91"

/lab host="DH10B"

/notc="vector: PCMV-SPORT6"

1..2538

/gene="SLA2"

/note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845, SLAP2"

/db\_xref="LocusID:84174"

/db\_xref="MIM:606577"

363..1148

/codon\_start=1

/product="Src-like adaptor 2, isoform a"

/protein\_id="AAH42041.1"

/db\_xref="GI:27469843"

/db\_xref="LocusID:84174"

/translation="WGSLSPPSRKSLPSPSSVQGGPVTMEARSKATAVALGSPFP AGGPALSLRLGLEPTIVSEGDWTVLSEVGREYNI PSYHAKVSHGWLVEGLSRE KAEELLLPGNPGGAFLRESQTRGSLSVLRSPASMDRINHYRIHCLDNGWLYI

## gene

## CDS

misc\_feature

SPRLTFPSLQALVDHYSELADDICLLKEPCVLQAGLPCKDIPLPVTVORTPLNWK ELDSLLPSEATGEESLSEGLRESUSFYISLNDEAVSLDDA"

468..632

/note="SH3; Region: SH3 domain. SH3 (src homology 3) domains are often indicative of a protein involved in signal transduction related to cytoskeletal organization. First described in the src cytoplasmic tyrosine kinase. The structure is a partly opened beta barrel"

/db\_xref="CDD:pfam00018"

642..890

/note="SH2; Region: SH2 domain"

/db\_xref="CDD:pfam00017"

misc\_feature

ORIGIN

Alignment Scores:

Pred. No.: 1,538-101 Length: 2538

Score: 1347.00 Matches: 260

Percent Similarity: 99.62% Conservatives: 0

Best Local Similarity: 99.62% Mismatches: 1

Query Match: 99.56% Indels: 0

DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x BC042041 (1-2538)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20

Db 363 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 422

QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40

Db 423 CAAGCCAGGGACCTGTACCATGGAAGCAGAGAGACCAAGCCACAGCCGTGCGCCCTG 482

QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGlyProLeuThr 60

Db 483 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCTGCTGAGACTCGGGAGCCATTGACC 542

QY 61 IleValSerGluAspGlyValSerThrPheValLeuSerGluValSerGlyArgGluTyr 80

Db 543 ATCGTCTCTGAGATGGAGACTGGTGACGGTGGTCTCTGAACTCTCAGGCAGAGAGTAT 602

QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100

Db 603 AACATCCCAGCGTCCAGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGCGCTGAGC 662

QY 101 ArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyValaPheLeuIle 120

Db 663 AGGAGAAAGCAGAGAACTGCTGTGTGTACCTGGAAACCTTGGAGGGCCCTTCTCATC 722

QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140

Db 723 CGGAGAGCCACACCCAGGAGAGGCTTACTCTCTGTCTCAGTCCGCTCAGCGCCCTGCA 782

QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160

Db 783 TCCTGGGACCGGATCAGACACTACAGGATCCACTCCCTTGACAAATGGCTGGCTGACATC 842

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180

Db 843 TCACCGGCGCTCACCTTCCCTCTCACTCCAGGCGCTGGTGACCATTTACTCTGAGCTGGC 902

QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200

Db 903 GATGACATCTGCTGCTACTCAAGGAGCGCTGTGCTTCGACAGAGGCGTGGCCGCTCCCT 962

QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220

Db 963 GGCAGGATATACCTCTACCTGTGACTGTGCAGAGGACACCACTCACTGGAAGAGCTG 1022

QY 221 AspSerSerLeuLeuPheSerGluAlaThrGlyGluGluSerLeuSerGluGly 240

Db 1023 GACAGTCCCTCTCTGTTTCTGAGCTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGGT 1082

QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260

Db 1083 CTCGGGAGTCCCTCAGCTTCTACATGAGCTGAATGACGAGGCTGCTCTTTGGATGAT 1142  
Qy 261 Ala 261  
Db 1143 GCC 1145  
RESULT 9  
AX452880  
LOCUS AX452880 2567 bp DNA linear PAT 06-JUL-2002  
DEFINITION Sequence 1 from Patent WO0242457.  
ACCESSION AX452880  
VERSION AX452880.1 GI:21712520  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villar, J.J. and Kanner, S.B.  
TITLE Cloning and expression of human slap-2: a novel sh2/sh3 domain-containing human slap homologue having immune cell-specific expression  
JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;  
Bristol-Myers Squibb Co. (US)  
FEATURES  
source  
1..2567  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.55e-101 Length: 2567  
Score: 1347.00 Matches: 260  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.56% Indels: 0  
DB: 6 Gaps: 0  
US-09-939-853A-75 (1-261) x AX452880 (1-2567)  
Qy 1 MetGlySerLeuProSerArgAlaGlySerLeuProSerProSerLeuSerSerVal 20  
Db 415 ATGGAGTCTGCCAGCAGAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 474  
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
Db 475 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 534  
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
Db 535 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGCGCTGAGACTCGGGAGCCATTGACC 594  
Qy 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
Db 595 ATCGTCTCTGAGGATGGAGACTGTGGACGCTGTGCTGAAGTCTCAGGCAGAGACTAT 654  
Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
Db 655 AACATCCCGCAGCTCCAGTGGCCCAAGTCTCCCATGGTGGCTGTATGAGGGCCTGAGC 714  
Qy 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120  
Db 715 AGGAGAAAGCAGAGGAATCTGTTTACTTGGGAACCTCGAGGGCCCTTCTCATC 774  
Qy 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
Db 775 CGGGAGAGCCAGCAGAGAGGCTTTACTCTCTGTGAGTCCGCCCTCAGCCGCCCTGCA 834  
Qy 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160

Db 835 TCCTGGGACCGGATCAGACACTACAGGATCCATGCTTGCATATGGCTGGCTGATCATC 894  
Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
Db 895 TCACGGCGCTTCACCTTCCCTCCTCAGGCCCTGGTGACCATTTACTCTGAGCTGGCG 954  
Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
Db 955 GATGACATCTGCTGCTTACTCAGAGGCCCTGTGTCTCGAGAGGCTGGCCGCTCCCT 1014  
Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
Db 1015 GGCAGGATATACCCCTACCTGTGACTGTGTCAGAGGACACCCACTCAATGGAAAGACTG 1074  
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
Db 1075 GACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTCTCAGTGGGT 1134  
Qy 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
Db 1135 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTCTTTGGATGAT 1194  
Qy 261 Ala 261  
Db 1195 GCC 1197  
RESULT 10  
AX780857  
LOCUS AX780857 2788 bp DNA linear PAT 14-JUL-2003  
DEFINITION Sequence 3014 from Patent WO03039443.  
ACCESSION AX780857  
VERSION AX780857.1 GI:32697851  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Haerlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S., Dugas, M., Ellis, R., Brose, B. and Mergenthaler, S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 3014 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE);  
Ludwig-Maximilians-Universitaet Muenchen (DE); Haerlach, Torsten,  
PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)  
FEATURES  
source  
1..2788  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.32e-101 Length: 2788  
Score: 1341.00 Matches: 259  
Percent Similarity: 99.23% Conservative: 0  
Best Local Similarity: 99.23% Mismatches: 2  
Query Match: 99.11% Indels: 0  
DB: 6 Gaps: 0  
US-09-939-853A-75 (1-261) x AX780857 (1-2788)  
Qy 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
Db 387 ATGGGAAGTCTCCCGCAGAGAGAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 446  
Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
Db 447 CAAGGCCAGGACCTGTGACCATGGAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 506  
Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyProLeuThr 60  
Db 507 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCTGAGACTCGGGAGCCATTGACC 566

QY 61 lleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 567 ATCTGCTCTGAGGATGGAGACTGTGTGACGGTGTCTGAAGTCTCAGGACAGAGATAT 626  
 QY 81 AsnIleProSerValHisValGlyValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 DB 627 AACATCCCGCGCTCCACGTCGCCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 686  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 DB 687 AGGAGAAAGCAGAGGAACCTCTGTGTACTCCGGAACCCCTGGAGGGGCTTCCTCATC 746  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 747 CGGAGAGCCAGACAGGAGGCTTACTCTCTGTCACTCCGCTCAGCGGCCCTGCA 806  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 DB 807 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCCCTTGACAAATGGCTGGCTGTACATC 866  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 DB 867 TCACCGCGCTCACCTTCCCTCACTCAGGCGCTGTGGACCACTTACTCTGAGCTGGCG 926  
 QY 181 AspAspIleCysCysLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 DB 927 GATGACATCTCTGCTACTCAAGAGCGCTGTCTCTCAGAGGGCTGGCGCCGCTCCCT 986  
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 DB 987 GGCAGAGATATACCCCTACCTGTACTGTGAGAGGACACCACTCAACTGGAAAGAGCTG 1046  
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
 DB 1047 GACAGCTCCCTCTGTTTCTGAACTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGCT 1106  
 QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
 DB 1107 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACAGGCTGTCTCTTTGGATGAT 1166  
 QY 261 Ala 261  
 DB 1167 GCC 1169  
 RESULT 11  
 AX511155  
 LOCUS  
 DEFINITION Sequence 6 from Patent WO0242452.  
 ACCESSION AX511155  
 VERSION AX511155.1 GI:23392047  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1  
 AUTHORS  
 TITLE Adapter gene  
 JOURNAL Patent: WO 0242452-A 6 30-MAY-2002;  
 The Hospital for Sick Children (CA)  
 FEATURES  
 Location/Qualifiers  
 1..737  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,426-90 Length: 737  
 Score: 1196.50 Matches: 240  
 Percent Similarity: 92.0% Conservative: 2  
 Best Local Similarity: 91.25% Mismatches: 2

Query Match: 88.43% Indels: 19  
 DB: 6 Gaps: 1  
 US-09-939-853a-75 (1-261) x AX511155 (1-737)  
 QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 DB 1 ATGGGAAGTCTCCCGCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAATCTCTGTC 60  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 DB 61 CAAGGCCAGGGACCTGTGTACCATGAAGCAGAGAGAAAGCAAGGCCACGCCCTGGCCCTG 120  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 DB 121 GGCAGTTTCCCGCGAGGTGGCCCGCGAGCTGTCTGCTGAGACTCGGGAGCCATTGACC 180  
 QY 61 lleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 181 ATCTGCTCTGAGGATGGAGACTGGTGGACGGTGTCTCTGAAGTCTCAGGCAGAGAGATAT 240  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 DB 241 AACATCCCGCAGCGTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCTGAGC 300  
 QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 DB 301 AGGAGAAAGCAGAGGAACCTGTGTGTACTCGGGAACCCCTGGAGGGGCTTCTCTCATC 360  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 361 CGGAGAGCCAGACACAGGAGGCTTACTCTCTGTCACTCCGCTCAGCCGCGCTGCA 420  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 DB 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTCCCTTGACAATGGCTGGCTGTACATC 480  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 DB 481 TCACCGCGCTCACCTTCCCTCACTCAGGCGCTGTGGACCATTC----- 528  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200  
 DB 529 -----TCTGAGGGCTGGCGCGCTCCC 549  
 QY 200 OGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220  
 DB 550 TGGCAAGGATATACCCCTACCTGTGACTGTGGGAGGACACCACTCACTGGAAAGAGCT 609  
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGl 240  
 DB 610 GGACAGCTCCCTCTGTTTCTGAACTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGG 669  
 QY 240 yLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA 260  
 DB 670 TCTCCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGAGCGAGGCTGTCTCTTTGGATG 729  
 QY 260 spAla 261  
 DB 730 ATGCC 734  
 RESULT 12  
 AF290986  
 LOCUS  
 DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,  
 complete cds; alternatively spliced.  
 ACCESSION AF290986  
 VERSION AF290986.1 GI:17351922  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 737)  
 Loreto, M.P. and McGlade, C.J.  
 Cloning and characterization of human Src-like adaptor protein 2  
 and a novel splice isoform, SLAP-2-v  
 Oncogene 22 (2), 266-273 (2003)  
 MEDLINE 22415750  
 PUBMED 12527895  
 REFERENCE 2 (bases 1 to 737)  
 Loreto, M.P. and McGlade, C.J.  
 Direct Submission  
 TITLE Brain Tumour Research Centre, Hospital for  
 Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada  
 JOURNAL Location/Qualifiers  
 FEATURES 1. 737  
 source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /tissue\_type="thymus"  
 1. 633  
 /note="SLAP-2-v"  
 /codon\_start=1  
 /product="Src-like adaptor protein-2 splice isoform"  
 /protein\_id="AAL38198.1"  
 /db\_xref="GI:17351923"  
 /translation="MGSLEPRKSLPSPSSVQGGPVTMEAKSKATAVALGSGFP  
 AGPAEISLRIGELTIVSDGWTVLSEVSGEYINPISVAKVSHGWLYEGLSRE  
 KAEELLIPNGGAFILIRSDTRRGYSLSRLRPASWDRIHRIHCLDNGWLYI  
 SPRLTFPSLQALVDHYSEGWPAPQWGTPTCDCAEDTTQLERAGQLPPVF"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9, 42e-90 Length: 737  
 Score: 1196.50 Matches: 240  
 Percent Similarity: 92.02% Conservative: 2  
 Best Local Similarity: 91.25% Mismatches: 2  
 Query Match: 88.43% Indels: 19  
 DB: 9 Gaps: 1  
 US-09-939-853A-75 (1-261) x AF290986 (1-737)  
 QY 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTGCCAAGCCCAAGCTGAGTCTCTGTC 60  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 61 CAAGGCCAGGGACCTGTGACCATGGAAGCAGAGAGAGAGGCGCACAGCGTGGCCCTG 120  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 181 ATCTGCTCTGAGGATGGAGACTGGTGGACGGTGTCTGAAGTCTCAGGAGAGAGAT 240  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
 Db 241 AACATCCCAGAGTCCAGCTGGCCAAAGTCTCCCATGGTGGTGGTGTATGAGGCGCTGAGC 300  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 Db 301 AGGGAGAAAGCAGAGAACTGT 360  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 361 CGGGAGAGCCAGACACAGAGAGAGGCTTCTACTCTCTGTCTGTCTGTCTGTCTGTCTGTCT 420  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 Db 421 TCTGGGACCGGATCAGACACTACAGATCCACTGCTTGTGCAATGGTGGCTGTGTATCATC 480

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 481 TCACCGCCCTCACCCTCCCTCTCACTCAGGCCCTGTGGACCATTAC----- 528  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPyr 200  
 Db 529 -----TCTGAGGGCTGGCCCGCTCCC 549  
 QY 200 cglLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220  
 Db 550 TGGCAGAGATATACCCCTACTGTGACTGTGGAGGACACACTCACTCACTGGAAGAGCT 609  
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluG 240  
 Db 610 GGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTCTCAGTGAGGG 669  
 QY 240 vLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA 260  
 Db 670 TCTCGGAGTCCCTCACTTCTACATCAGCTGATGAGCGAGCTCTCTCTTGGATG 729  
 QY 260 spAla 261  
 Db 730 ATGCC 734  
 RESULT 13  
 AX511151  
 LOCUS Sequence 2 from Patent WO0242452.  
 DEFINITION AX511151  
 ACCESSION AX511151  
 VERSION AX511151.1 GI:23392045  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS McGlade, J.C. and Loreto, M.P.  
 TITLE Adapter gene  
 JOURNAL Patent: WO 0242452-A 2 30-MAY-2002;  
 The Hospital for Sick Children (CA)  
 FEATURES  
 Location/Qualifiers  
 1. 777  
 /organism="Mus musculus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:10090"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7, 96e-76 Length: 777  
 Score: 1028.00 Matches: 208  
 Percent Similarity: 85.50% Conservative: 16  
 Best Local Similarity: 79.39% Mismatches: 34  
 Query Match: 75.98% Indels: 4  
 DB: 6 Gaps: 3  
 US-09-939-853A-75 (1-261) x AX511151 (1-777)  
 QY 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 1 ATGGGAAGTGTTCAGCAGAGAGGAAACCC---TCCAGGCCCCAGCCCTCTCTGTGT 57  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 58 CCAGACCCAGAACCCGTGTCCATGCGACCCAGAAAGACACAGAGTACAGCTGTGGCCCTG 117  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 118 GGCAGTTTCCAGCAGGTGAAACAGGCGCAGACTATCTCTGAGACTTCGGGGAGCGCTGACC 177  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 178 ATCATCTCTGAGGATGGAGATTGGTGACAGTCCACTCGAAGTCTCAGGCGAGAGATAC 237

```

QY      51  AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
      238  CACATGCCAGTGTGTATGTGCTAAAGTCGCCACCGGCTGGCTGTACGAGGGGCTGAGC 297
QY      101  ArgGluLysAlaGluGluLeuLeuLeuProGlyValAsnProGlyGlyAlaPheLeuLeu 120
      258  CGGAGAAAGCGGAGGAACCTACTCTGTACTCTGGAAACCCCGGAGGGGCTTCTCTCATC 357
QY      121  ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
      358  CGGAGAGCCAGACCCAGAGAGGCTGTATTCCTGTCCGTCCGACTCAGCGGCCCTGCA 417
QY      141  SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
      418  TCTTGGGACCGATCAGACACTACAGGATCAGCGCTTGTGCAATGGCTGGCTGTATCATC 477
QY      161  SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
      478  TCACCTCGCCTCACCTTCCCTCACTCCACGCTTGGTGGAGCATTACTCTGAGCTAGCA 537
QY      181  AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
      538  GATGGCATCTGTCTGCCCTCAGGAGCGGTGTCTGCGAAGAGCTTGGGCCACTACCT 597
QY      201  GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLeu 220
      598  GCGAAGATACACCTCCACCTGTGACTGTGCCAATCATCATCAATAATGGAAAAGCTG 657
QY      221  AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuSerGlu 239
      658  GACCCGAGCCTCTCTTTCTGGAACACCTCGGAGTGGGGAGGCATCTCTCTCAGTGAG 717
QY      240  GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
      718  GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCTGGCTGAGGAC-----CCCTTGGAT 771
QY      260  AspAla 261
      772  GATGCT 777

RESULT 14
LOCUS   AX511150
DEFINITION Sequence 1 from Patent WO0242452.
ACCESSION AX511150
VERSION   AX511150.1 GI:23392044
KEYWORDS
SOURCE   Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS  McGlade,J.C. and Loreto,M.P.
TITLE    Adapter gene
JOURNAL  Patent: WO 0242452-A 1 30-MAY-2002;
        The Hospital for Sick Children (CA)
FEATURES
source   1..1348
         /organism="Mus musculus"
         /mol_type="unassigned DNA"
         /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 1,51e-75 Length: 1348
Score: 1028.00 Matches: 208
Percent Similarity: 85.50% Conservative: 16
Best Local Similarity: 79.39% Mismatches: 34
Query Match: 75.98% Indels: 4
DB: 6 Gaps: 3

US-09-939-853A-75 (1-261) x AX511150 (1-1348)

```

```

QY      1  MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal 20
      282  ATGGGAAGTTTGTCCAGCAGAGGGAACCC---TCCAGCCCCAGCCCCAGCTCCTCTGGT 338
QY      21  GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
      339  CCAGACCAGAACCGGTGTCATGTCAACACGAAAACACACAGGTTCACAGCTGTGGCCCTG 398
QY      41  GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
      399  GGCAGTTTCCCGACAGGTGACAGGCCAGACTATCTCTGAGACTCGGGAGGCGCTGACC 458
QY      61  IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGluTyr 80
      459  ATCATCTCTGAGATGAGATTGGTGGACAGTCCAGTCCGAAAGTCTCAGGCAGAGATAC 518
QY      81  AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
      519  CACATGCCAGTGTGTATGTGGCTAAAGTCGCCACCGGTGGCTGTACGAGGCCCTGAGC 578
QY      101  ArgGluLysAlaGluGluLeuLeuLeuProGlyValAsnProGlyGlyAlaPheLeuLeu 120
      579  CGGAGAAAGCGGAGGAACCTACTCTCTGTACTCTGGAAACCCCGGAGGGGCTTCTCTCATC 638
QY      121  ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
      639  CGGAGAGCCAGACCCAGAGAGGCTGTATTCCTGTCCGTCCGACTCAGCGGCCCTGCA 598
QY      141  SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
      699  TCTTGGGACCGATCAGACACTACAGGATCAGCGCTTGTGCAATGGCTGGCTGTATCATC 758
QY      161  SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
      759  TCACCTCGCCTCACCTTCCCTCACTCCACGCTTGGTGGAGCATTACTCTGAGCTAGCA 818
QY      181  AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
      819  GATGGCATCTGTCTGCCCTCAGGAGCGGTGTCTCTCAGAAAGCTTGGGCCACTACCT 878
QY      201  GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLeu 220
      879  GCGAAGATACACCTCCACCTGTGACTGTGCCAATCATCATCAATAATGGAAAAGCTG 938
QY      221  AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuSerGlu 239
      939  GACCCGAGCCTCTCTTTCTGGAAGCACCTCGGAGTGGGGAGGCATCTCTCTCAGTGAG 998
QY      240  GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
      999  GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCTGGCTGAGGAC-----CCCTTGGAT 1052
QY      260  AspAla 261
      1053  GATGCT 1058

RESULT 15
LOCUS   AF287467
DEFINITION Mus musculus Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF287467
VERSION   AF287467.1 GI:17351918
KEYWORDS
SOURCE   Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS  Loreto,M.P., Berry,D.M. and McGlade,C.J.
TITLE    Functional cooperation between c-Cbl and Src-like adaptor protein 2
        in the negative regulation of T-cell receptor signaling
JOURNAL  Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
MEDLINE  22022020

```

PUBMED 12024036  
REFERENCE 2 (bases 1 to 1348)  
AUTHORS Loreto, M.P. and McGlade, C.J.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada  
FEATURES  
source Location/Qualifiers  
1..1348  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/dev\_stage="day 15 embryo"  
282..1061  
/feature="SLAP-2"  
/feature="SLAP-2"  
/codon\_start=1  
/product="Src-like adaptor protein-2"  
/protein\_id="AAL38196.1"  
/db\_xref="GI:1731919"  
/translation="MGSLSRGKTSPPSSGPDPEVSMQPERHKVTAVALGSPFA  
GEQARLSLRIGEPITIIISERQDMWTVQSEVSGREVHMPVYVAVKAVHGLYGLSREK  
AEELLLIPGPGNPLIRSQTRGCVSLVRLSRPASNDRIHRYRIQLDNGWLYIS  
PRITFPLHALVEHYSLADGICPLREPCVLOKGLPGLPOTPPPTVITSSLNWKK  
LDRLDFLEAPASGEANLLSEGLRESLSYISLAEDPLDDA"  
381..551  
/misc\_feature /note="Region: SH3 domain"  
552..824  
/misc\_feature /note="Region: SH2 domain"  
1301..1306  
/polyA\_signal  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,51e-75 Length: 1348  
Score: 1028.00 Matches: 208  
Percent Similarity: 85.50% Conservativeness: 16  
Best Local Similarity: 79.39% Mismatches: 34  
Query Match: 75.98% Indels: 4  
DB: 10 Gaps: 3  
US-09-939-853A-75 (1-261) x AF287467 (1-1348)  
QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
DB 282 ATGGAGAGTTTTCAGCAGAGAGGAAACCC---TCCAGCCCCCAGCTCTCTGGT 338  
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
DB 339 CCAGACCAGGAACCGCTGTCATGCACACCAAGAGACACACAGGTCACAGCTGTGGCCCTG 398  
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
DB 399 GGCAGTTCAGCAGAGGAGGAGCAGCAGCTATCTGTGAGACTCCGGGAGCCGCTGACC 458  
QY 61 IleValSerGluAspGlyAspTyrTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
DB 459 ATCATCTCTAGGATGGAGATTGGTGACAGTCCAGTCCGAAAGTCTCAGGCAGAGAGTAC 518  
QY 81 AsnIleProSerValHisValGlyValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
DB 519 CACATGCCACGTGTGTATGTGGTAAAGTCGCCACCGGTGGCTGTACGAGGGCCCTGAGC 578  
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120  
DB 579 CGGAGAAAGCCGAGGAACCTACTCTGTACTGGGAACCCCGGAGGGGCTTCTCTATC 638  
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
DB 639 CGGGAGAGCCAGACAGAGAGGCTGTATTCCCTGTCCGTCCTGAGCTCAGCCGCCCTGCA 698  
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
DB 699 TCTTGGGACCGGATCAGACACTACAGGATACAGCGCTTGTGACATGGCTGGCTATCATC 758  
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180

DB 759 TCACCTCGCCTCACCTTCCCTCCTCACCTTGGTGGAGCATTACTCTGAGCTAGCA 818  
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
DB 819 GATGCATCTGCTGTCCCTCAGGAGCCGCTGTCTCTGCAAGACTTGGGCCACTACCT 878  
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
DB 879 GGCAAGATACACCTCCACCTGTGCTGTGCCAACATCATCATCAATTTGAAAAGCTG 938  
QY 221 AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuSerGlu 239  
DB 939 GACCCAGCCTTCTTCTGAGAGCCTTGGAGTGGGAGGAGCATCTCTCTCAGTGAG 998  
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259  
DB 999 GGGCTCCGAGAGTCCCTCAGTTCCTATCATCAGCCTGGCTGAGGAC-----CCCTTGGAT 1052  
QY 260 AspAla 261  
DB 1053 GATGCT 1058  
Search completed: February 20, 2004, 10:48:33  
Job time : 3102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:33:07 ; Search time 368 Seconds  
(without alignments)  
3012.988 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSIPSRKSLPESLSVV.....RESLSPVLSINDEAVSLDDA 261

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO.spool.p/US09939853/runat.19022004.145338.24420/app.query.fasta\_1.455  
-DB=N Geneseq 29Jan04 -OPT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LCOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human.cod  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09939853.ecgn.1.708.0/runat.19022004.145338.24420 -NCFU=6 -ICFU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002s: \*  
7: Geneseqn2003as: \*  
8: Geneseqn2003bs: \*  
9: Geneseqn2003cs: \*  
10: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1353	100.0	1183	6	ABK61465 Human CDN
2	1347	99.6	786	6	AAL44089 Human ORF
3	1347	99.6	786	6	ABQ74343 Human Src
4	1347	99.6	2567	6	AAD43980 Human Src
5	1269	93.8	837	3	AAC77202 Human ORF
6	1206.5	89.2	1413	6	ABQ99374 Human cod
7	1196.5	88.4	737	6	AAL44090 Mouse MAR
8	1128.5	83.4	2049	5	AAS74750 DNA encod

9	1028	76.0	1348	6	AAL44087 Mouse mod
10	826	61.0	763	5	ABQ98670 Human ORF
11	731	54.0	603	5	AAS74748 DNA encod
12	586	43.3	875	6	ABQ99151 Human ORF
13	488	36.1	3756	7	ABX62975 Human act
14	487	36.0	2109	4	AAS02049 DNA encod
15	487	36.0	2665	6	ABL65189 Lung can
16	487	36.0	2665	6	ABK83738 Human CDN
17	487	36.0	2665	7	ACC81091 Human Src
18	452.5	33.4	444	6	ABQ98669 Human ORF
19	370.5	27.4	2298	6	ABK83935 Human CDN
20	370.5	27.4	2298	9	ADD19015 Human dis
21	360.5	26.6	1926	6	ABK83940 Human CDN
22	360.5	26.6	2015	6	ABL66673 Lung can
23	360.5	26.6	2015	6	ABK83939 Human CDN
24	360.5	26.6	2015	9	AAD62155 Human hae
25	360.5	26.6	2341	9	ABE53793 Human pro
26	358.5	26.5	1911	6	ABK63704 Rat sequ
27	358.5	26.5	1911	9	ADB58205 Toxicity-
28	358.5	26.5	1911	9	ADB52725 Primary r
29	348	25.7	1530	8	ADA02968 Human LCK
30	348	25.7	1530	9	ADB72706 Human LCK
31	348	25.7	1530	9	ADC85448 Human LCK
32	348	25.7	2032	3	AZA46491 PKA subst
33	348	25.7	2032	7	ACC72850 Human can
34	348	25.7	2032	8	ADA02967 Human LCK
35	348	25.7	2032	9	ADB72705 Human LCK
36	348	25.7	2032	9	ADC85447 Human LCK
37	348	25.7	2032	9	ADB72705 Human LCK
38	340	25.1	2034	6	ABE40448 Human pro
39	340	25.1	2129	7	ABG65202 CDNA enco
40	340	25.1	2129	7	ACA56854 Human sig
41	340	25.1	2129	7	ACC81082 Human t-t
42	338.5	25.0	1530	8	ACA64888 Human p56
43	338.5	25.0	1530	9	ADA02965 Mouse Lck
44	338.5	25.0	1530	9	ADB72703 Mouse Lck
45	338.5	25.0	1530	9	ADC85445 Mouse Lck
			2100	8	ADA02964 Mouse Lck

ALIGNMENTS

RESULT 1

ABK61465  
ID ABK61465 standard; cDNA; 1183 BP.

XX AC ABK61465;

XX DT 18-JUN-2002 (first entry)

XX DE Human cDNA encoding protein NOV13.

XX KW Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;

KW cell signal processing disorder; metabolic pathway modulation disorder;

KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;

KW uterus cancer; immune response; graft-versus-host disease;

KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;

KW hypertension; congenital heart defects; multiple sclerosis; inflammation;

XX KW Albright hereditary osteodystrophy.

XX OS Homo sapiens.

XX FN WO200216599-A2.

XX PD 28-FEB-2002.

XX PF 27-AUG-2001; 2001WO-US026510.

XX PR 25-AUG-2000; 2000US-0228191P.

XX PR 08-FEB-2001; 2001US-0267300P.

XX PR 20-FEB-2001; 2001US-0269961P.

XX PR 20-MAR-2001; 2001US-0277337P.

XX PA (CURA-) CURAGEN CORP.

(CORT-) COR THERAPEUTICS INC.

PA Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shinkets RA;  
 PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;  
 XX WPI; 2002-280937/32.  
 DR P-PSDB; AAU91308.

XX New polypeptides for treating or preventing a disorder associated with  
 PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.  
 PT

PS Claim 1; Page 98; 263pp; English.

XX The invention relates to an isolated polypeptide (NOVX) a mature form of  
 CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide  
 CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
 CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it,  
 CC and antibody against it, are useful for treating or preventing (e.g. by  
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOVX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation or Albrit  
 CC hereditary osteodystrophy and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC for developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence encodes a NOVX protein

XX Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,41e-109 Length: 1183  
 Score: 1353.00 Matches: 261  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x ABK61465 (1-1183)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
 DB 398 ATGGAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCCTCTGTC 457  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 DB 458 CAAAGCCAGGACCTGTGACCATGTGAGCAGAGAGACAGGACAGCGGTGGCCCTG 517  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgGlyGluProLeuThr 60  
 DB 518 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCTGAGACTCGGGAGCCATTGACC 577  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 578 ATCGTCTCTGAGATGGAGACTGGTGGACGGTGTCTGTAAGTCTCAGGAGAGAGAT 637  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 DB 638 AACATCCCGAGCTCCACGTGGGCAAGTCTCCCATGGTGGCTGTATGAGGGCTCAGC 697  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 DB 698 AGGGAGAAAGCAGAGAACTGCTGTGTGTACCTGGGAAACCTTGGAGGGGCTTCTTCATC 757

QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 758 CGGAGAGCCAGACCCAGAGAGGCTCTTACTCTCTGTCTGCTCAGCCCTCAGCGCCCTGCA 817  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAspGlyTrpLeuTyrIle 160  
 DB 818 TCCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGCATATGGCTGGCTGTATC 877  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 DB 878 TCACCGCGCTCACCTCCCTCCTCCTCAGGCCCTGGTGGACCATTTACTCTGAGCTGGCG 937  
 QY 191 AspAspIleCysCysLeuLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 DB 938 GATGACATCTGCTCTCTCTCAGGAGCCCTGTGTCTCTCAGAGGGCTGGCCCTCCCT 997  
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 DB 998 GGCAGGATATACCCCTACCTGTGACTGTGAGAGGACACCCTCAACTGGAAAGAGCTG 1057  
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGluGly 240  
 DB 1058 GACAGCTCCCTCTCTTCTGAAGTGCACAGGGGAGGAGTCTCTCTCAGTGAGGAT 1117  
 QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
 DB 1118 CTCGGGAGTCCCTCAGCTTCTATCATCAGCTGAATGACGAGGCTGTCTCTTGGATGAT 1177  
 QY 261 Ala 261  
 DB 1178 GCC 1180  
 RESULT 2  
 AAL44089  
 ID AAL44089 standard; cDNA; 786 BP.  
 XX AAL44089;  
 XX 03-OCT-2002 (first entry)  
 XX Human modulator of antigen receptor signalling protein coding sequence.  
 DE Human; gene; ss; gene therapy; modulator of antigen receptor signalling;  
 KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP; disorder;  
 KW myeloid malignancy; acute myelogenous leukemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 CDS 1..786  
 FT /\*tag= a  
 FT /product= "Human MARS protein"  
 FT  
 XX WO200242452-A2.  
 XX 30-MAY-2002.  
 XX 26-NOV-2001; 2001WO-CA001662.  
 XX 27-NOV-2000; 2000CA-02324663.  
 XX (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 XX Mcglade JC, Loreto MP;  
 XX WPI; 2002-566564/60.  
 DR P-PSDB; AA015457.  
 XX New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX

PS Claim 12; Page 75; 110pp; English.

XX The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present cDNA sequence encodes a human MARS protein  
 XX

SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.79e-109 Length: 786  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.56% Indels: 0  
 DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x AAL44089 (1-786)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
 DB 1 ATGGGAAGTCTGCCAGCAGAGAATAATCTCTGCAAGCCCAAGCTTGGTCTCTGTC 60  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40  
 DB 61 CAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCAGCGCTGGCCCTG 120  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluSerLeuArgGlyGluProLeuThr 60  
 DB 121 GGCAGTTTCCCGCAGGTGGCCGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 181 ATCTCTCTGAGGATGGAGACTGTGGACGGTGTCTGTGAAGTCTCAGGAGAGAGTAT 240  
 QY 81 AsnIleProSerValHisValGlyValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
 DB 241 AACATCCACCGTCCAGCTGGCCAAAGTCTCCATGGGTGGCTGTATGAGGCGCTGAGC 300  
 QY 101 ArgGluIysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLe 120  
 DB 301 AGGAGAGAAGCAGAGAACTGCTGTGTACCTGGGAACCCCTGGAGGGGGCTTCTCTCATC 360  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 361 CGGGAGAGCCAGACAGAGAGAGGCTTTACTCTCTGTCAGTCCGCTCAGCCGCGCTGCA 420  
 QY 141 SerTipAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160  
 DB 421 TCTTGGGACCGGATCAGACACTACAGGATCCACTGCCTTGACAATGGCTGGCTGTACATC 480  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 DB 481 TCACCGCCCTCACTCCCTTCCCTCACTCCAGGCGCTGGTGGACCAATTACTCTGAGCTGGCG 540  
 QY 181 AspAspIleCysCysLeuLeuIysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 DB 541 GATGACATCTGCTGCTACTCAAGGAGCGCTGTGCTCTGCAGAGGCTGGCCGCTCCCT 600  
 QY 201 GlyIysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpIysGluLeu 220  
 DB 601 GGCAAGGATATACCCCTCACTGTGACTGTGCAGAGGACACCACTCAACTGGAAGAGCTG 660  
 QY 221 AspSerSerLeuLeuPheSerGluAlaIleThrGlyGluGluSerLeuLeuSerGluGly 240  
 DB 661 GACAGCTCCCTCTCTTTCTGAAGCTGCCACAGGGAGAGTCTCTCTCAGTGAGGCT 720  
 QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260

DB 721 CTCGCGGAGTCCCTCAGCTTCTATCATGAGCTGATGACGAGGCTGTCTCTTGGATGAT 780  
 QY 261 Ala 261  
 DB 781 GCC 783  
 RESULT 3  
 ABQ74343  
 ID ABQ74343 standard; cDNA; 786 BP.  
 XX AC ABQ74343;  
 XX 15-OCT-2002 (first entry)  
 XX Human Src-like inhibitory molecule (SLIM) encoding cDNA.  
 DE Human; Src-like inhibitory molecule; SLIM; Src-like adaptor protein;  
 XX SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;  
 KW modulator; lymphocyte; Cbl; Gene therapy; immunodeficiency disorder;  
 KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;  
 KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;  
 KW Gene; ss.  
 XX Homo sapiens.  
 OS Key Location/Qualifiers  
 XX 1..786  
 FT CDS  
 FT /\*tag= a "SLIM"  
 FT /product= "SLIM"  
 FT /note= "Src-like inhibitory molecule"  
 XX WO200255707-A2.  
 XX 18-JUL-2002.  
 XX 10-JAN-2002; 2002WO-US000718.  
 XX 10-JAN-2001; 2001US-0260953P.  
 XX (RIGE-) RIGEL PHARM INC.  
 XX Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;  
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;  
 XX WPI; 2002-575432/61.  
 DR P-PSDB; ABP52187.  
 XX New src-like inhibitory molecule protein, useful for treating  
 PT immunodeficiency disorders and inflammatory disorders, comprises N-  
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.  
 XX Claim 3; Fig 2A; 91pp; English.  
 XX The present sequence encodes the human Src-like inhibitory molecule  
 CC (SLIM) protein (I). The present invention describes a SLIM protein  
 CC comprising an N-terminal myristylation sequence, an N-terminal SH2  
 CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising  
 CC an N-terminal myristylation sequence and an N-terminal SH2 domain which  
 CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and  
 CC anti-HIV activities, and can be used as a modulator of lymphocyte  
 CC activation, and of ubiquitination of a Cbl target protein, and in gene  
 CC therapy. (I) is useful for screening a bioactive agent capable of  
 CC to SLIM. (I) is also useful for screening a bioactive agent capable of  
 CC modulating SLIM binding. (I) or its fragments is useful in the study or  
 CC in the treatment of conditions which involves this function or  
 CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or  
 CC prevent SLIM associated disorders. (I) or the polynucleotide encoding it  
 CC (II) is useful for modulating leukocyte and/or platelet activation, for  
 CC modulating antigen receptor-induced signalling and activation in  
 CC leukocyte and/or platelets and for modulating antigen receptor-induced  
 CC signalling and activation in lymphocytes and/or mast cells. (I) or (II)  
 CC is also useful for modulating the basal activity of lymphocytes. (I) or

CC (II) is useful in the treatment of immunodeficiency disorders, such as  
 CC acquired immunodeficiency syndrome (AIDS), for the prevention and  
 CC treatment of acute inflammatory disorders, chronic inflammatory  
 CC disorders, autoimmune disorder and transplant rejection  
 XX  
 SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4, 79e-109 Length: 786  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 99.62% Conservatives: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.56% Indels: 0  
 DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x ABQ74343 (1-786)

Qy 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 Db 1 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGACTCTCTGTC 60

Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 61 CAAGCCAGGAGCTGTGACCATGGAGCAGAGAGAGAGCCAGCCGCTGGCCCTG 120

Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 121 GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTCGTGAGACTCGGGGAGCCATTGACC 180

Qy 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 181 ATCTCTCTGTAGGATGGAGACTGTGTGACGGTGTCTGAAGTCTCAGGCACAGAGTAT 240

Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 Db 241 AACATCCCGCAGCTCCAGTGGCCGACAGTCTCCATGGTGGTGTATGAGGCCCTGAGC 300

Qy 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 Db 301 AGGAGAAAGCAGAGGAACTCTCTTTTACCTGGGAACCTGGAGGGGCTTCTCCATC 360

Qy 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 361 CGGAGAGCCAGACACAGAGAGGCTCTTACTCTCTGTCAGTCCGCTCAGCCGCTGCA 420

Qy 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspIleGlyTrpLeuTyrIle 160  
 Db 421 TCCTGGGACCGATCAGACACTACAGGATCCACTGCTTGCACATGGCTGGCTGTACATC 480

Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 481 TCACCGCGCTCCTCCCTCCCTCCTCCTCAGCCCTGTGTGACCATTTACTGTAGCTGGCG 540

Qy 181 AspAspIleCysCysLeuLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 Db 541 GATGACACTGTGCTGCTACTCAAGGAGCCCTGTCTCTGACAGGGGTGGCGGCTCCCT 600

Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 Db 601 GCAAGGATATACCCCTACCTGTGACGTGACAGAGGACACCACTCAACTGAAAGAGCTG 660

Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
 Db 661 GACAGCTCCCTCTCTTTTCTGAGCTGCCACAGGGAGGAGTCTCTTCTCAGTGAGGCT 720

Qy 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
 Db 721 CTCGGGAGTCCCTCAGCTTCTATCATGACGCTGAATGACAGGGCTGTCTTCTTGGATAT 780

Qy 261 Ala 261  
 Db 781 GCC 783

RESULT 4  
 AAD43980  
 ID AAD43980 standard; cDNA; 2567 BP.  
 XX  
 AC AAD43980;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 DE Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.  
 XX  
 KW Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 415..1200  
 FT /\*tag= a  
 FT /product= "Human SLAP-2"  
 XX  
 PN WO200242457-A1.  
 XX  
 PD 30-MAY-2002.  
 XX  
 PF 20-NOV-2001; 2001WO-US043367.  
 XX  
 PR 22-NOV-2000; 2000US-0252545P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;  
 XX  
 DR WPI: 2002-463632/49.  
 XX  
 DR P-PSDB; AAE26357.  
 XX  
 PT Novel substantially purified human SH2/SH3-domain-containing adapter  
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic  
 PT intervention in immunological and inflammatory disorders and cancer.  
 XX  
 PS Claim 2; Fig 1; 85pp; English.  
 XX  
 CC The invention relates to a substantially purified human SH2/SH3-domain-  
 CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-  
 CC 2). The invention is useful for treating an immune disorder involving  
 CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is  
 CC useful for screening for antagonists or inhibitors of the interaction of  
 CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or  
 CC preventing diseases or disorders associated with aberrant or uncontrolled  
 CC cellular signal transduction, for determining those cellular signalling  
 CC molecules which associate with hSLAP-2 and which provide critical signals  
 CC for cell activation, and as effectors in methods to affect T- cell  
 CC activation. The invention is useful in screening assays to identify and  
 CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for  
 CC potential use to treat autoimmune diseases which may be caused by  
 CC hyperactivated B cells, as well as to treat diseases which may be caused  
 CC by hyperactivated T cells, in addition to other immune system related  
 CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,  
 CC inflammation disorders, diseases and conditions, rheumatoid arthritis,  
 CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's  
 CC and ulcerative colitis), allergies, particularly those involving  
 CC hyperactivity of B-cells and T- cells, or other immune cells, such as  
 CC mast cells or eosinophils, autoimmune diseases such as systemic lupus  
 CC erythematosus and multiple sclerosis, pulmonary diseases including  
 CC asthma, acute respiratory distress syndrome, and chronic obstructive  
 CC pulmonary disorder, tissue/ organ rejection and cancer. The invention is  
 CC useful in gene therapy. The present sequence is human SLAP-2 cDNA  
 XX  
 SQ Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.19e-108 Length: 2567  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.56% Indels: 0  
 DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x AAD43980 (1-2567)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 DB 415 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTCAGTTCCTCTGTC 474  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 DB 475 CAAGGCCAGGACCTGTGACCATGTAACACAGAGAGAACAGGCCACAGCCGCTGGCCCTG 534  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 DB 535 GGCAGTTTCCCGCAGGTGGCCCGCCAGCTGTCTGAGACTCGGGAGCCATTGACC 594  
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 595 ATCGTCTCTGAGGATGAGACTGTGTGGACGGTGTCTGTAAGTCTCAGGCAGAGAGTAT 654  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 DB 655 AACATCCCGAGCTCCAGTGGCCAAAGTCTCCCATGTTGGTGGTGTATGAGGGCTGAGC 714  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuLe 120  
 DB 715 AGGGAGAAAGCAGAGAGAACTGTGTGTGTACTCTGGGAACCTGGAGGGGCTTCTCTCATC 774  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 775 CGGGAGAGCCAGCAGGAGAGGCTTACTCTGTGTCAGTCCGCTCAGCCGCTCGCA 834  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 DB 835 TCCTGGGACCGATCAGACACTACAGGATCCACTGCCCTTGACATGGCTGTGTACATC 894  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 DB 895 TCACCGCGCCTCACCTTCCCTCACTCAGGAGCCCTGTGCTGCAGAGGGCTGGCCGCTGCG 954  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 DB 955 GATGACATCTGTGCTACTCAGGAGCCCTGTGCTGCAGAGGGCTGGCCGCTCCT 1014  
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 DB 1015 GCGAAGGATATACCCCTACTCTGACTGTGCAGAGGACACCACTCAACTGGAAGAGCTG 1074  
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
 DB 1075 GACAGCTCCCTCTGTTTCTGAACTGCCACAGGGAGAGTCTCTCTCTCAGTGAGGGT 1134  
 QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260  
 DB 1135 CTCGGGAGTCCCTCAGCTTCTACATCAGCCTGAATGACGAGCTGTCTCTTTGGATGAT 1194  
 QY 261 Ala 261  
 DB 1195 GCC 1197

RESULT 5

AAC77202  
 ID AAC77202 standard; cDNA; 837 BP.  
 XX  
 AC AAC77202;  
 XX

Alignment Scores:

Pred. No.: 3.67e-102 Length: 837  
 Score: 1269.00 Matches: 244  
 Percent Similarity: 99.19% Conservative: 1  
 Best Local Similarity: 98.79% Mismatches: 2

Sequence 837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;

08-FEB-2001 (first entry)

Human ORFX ORF2757 polynucleotide sequence SEQ ID NO:5513.

Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SIDA; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.

Homo sapiens.

WC200058473-A2.

05-OCT-2000.

31-MAR-2000; 2000WO-US008621.

31-MAR-1999; 99US-0127607P.

02-APR-1999; 99US-0127636P.

03-APR-1999; 99US-0127728P.

30-MAR-2000; 2000US-00540763.

(CURA-) CURAGEN CORP.

Shimkets RA, Leach M;

WPI; 2000-602362/57.

P-PSDB; AAB42993.

Novel nucleic acids and peptides derived from open reading frame X,  
 useful for treating e.g. cancers, proliferative disorders,  
 neurodegenerative disorders and cardiovascular disease.

Claim 5; Page 4692-4693; 5507pp; English.

AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 KW antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
 KW anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 KW cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 KW dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 KW antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The  
 sequences can be used for determining the presence of or predisposition  
 to, or preventing or treating pathological conditions associated with an  
 ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 proteins in gene therapy vectors. The proteins and nucleic acids may be  
 used to treat cancers, proliferative disorders, neurodegenerative  
 KW disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 KW diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 KW storage, systemic lupus erythematosus, severe combined immunodeficiency  
 KW (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 KW disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 KW cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 KW enhance coagulation; to inhibit thrombosis; and as a contraceptive

Query Match: 93.79% Indels: 0  
DB: 3 Gaps: 0

US-09-939-853A-75 (1-261) x AAC77202 (1-837)

QY 15 SerLeuSerSerValGlnGlyGlnGlyProValThrMetGluAlaGluArgSerLys 34  
DB 4 AGCTTGAGTTCCTCTGTCCAAAGCCAGGACCTGTGACCATGAGCAGAGAACCAAG 63

QY 35 AlaThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArg 54  
DB 64 GCCACAGCCGCGCCCTGGGCGAGTTCCCGCAGGTGGCCCGCGAGCTGTGCTGAGA 123

QY 55 LeuGlyGluProLeuThrIleValSerGluAspIleAspTyrThrValLeuSerGlu 74  
DB 124 CTCGGGAGCCATTGACCATCTCTGAGAGTGGAGACTGGTGAGCGGTCTGTGAA 183

QY 75 ValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysValSerHisGlyTyr 94  
DB 184 GTCTCAGGCAGAGATTAACATCCCGAGCGTCCACGTGGCCAAAGTCTCCCATGGTGG 243

QY 95 LeuTyrGluGlyLeuSerArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnPro 114  
DB 244 CTGTATAGGGCCCTGAGCAGGAGAAAGCAGAGACCTGTGTGTACCTGGGAACCT 303

QY 115 GlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTyrSerLeuSerVal 134  
DB 304 GGAGGGGCTTCTCATCCGGAGAGCCAGCAGGAGAGGCTTACTCTCTGTGAGTC 363

QY 135 ArgLeuSerArgProAlaSerTyrAspArgIleArgHisTyrArgIleHisCysLeuAsp 154  
DB 364 CGCTCAGCCGCTGCTATCTGGGACCGATCAGACACTACAGGATCCACTGCCCTTGC 423

QY 155 AsnGlyTyrLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAsp 174  
DB 424 AATGGCTGGCTGTACATCTACCGCGCTACCTTCCCTCACTCCAGGCCCTGGTGGAC 483

QY 175 HisTyrSerGluLeuAlaAspAspIleCysLeuLeuLysGluProCysValLeuGln 194  
DB 484 CATTACTGTAGCTGGCGGATGATCATCTGCTCCCTACTCAAGGAGCCCTGTGCTGCAG 543

QY 195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPro 214  
DB 544 AGGCTGGCCGCTCCCTGGCAAGATATACCTTACTGTGACTGTGAGAGGACACCA 603

QY 215 LeuAsnThrLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGlu 234  
DB 604 CTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGAG 663

QY 235 SerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu 254  
DB 664 TCTCTTCTCAGTGGGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCCTGAATGACGAG 723

QY 255 AlaValSerLeuAspAspAla 261  
DB 724 GCTGTCTCTTTGGATGATGCC 744

RESULT 6  
ABQ99374  
ID ABQ99374 standard; cDNA; 1413 BP.  
AC AC  
XX ABQ99374;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human coding sequence SEQ ID 107.  
XX  
KW Human; expressed sequence tag; EST; chromosome 20;  
KW haematopoietic disorder; central nervous system disease; viral infection;  
KW peripheral nervous system disease; non-healing wound; infectious disease;  
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;  
KW anti-allergic; anti-inflammatory; immunosuppressive; neuroprotective;

KW cytostatic; haemostatic; virucide; antibacterial; fungicide;  
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.  
XX Homo sapiens.  
XX WO200259260-A2.  
XX PD 01-AUG-2002.  
XX PF 16-NOV-2001; 2001WO-US042950.  
XX PR 17-NOV-2000; 2000US-00714936.  
XX (HYSE-) HYSEQ INC.  
XX PA Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX WPI; 2002-590824/63.  
XX DR N-PSDB; ABP64788.  
XX PT New isolated polynucleotide, useful in research, diagnostic or  
PT therapeutic methods, e.g. preventing or treating disorders involving  
PT aberrant protein expression or biological activity.  
XX  
PS Claim 1; SEQ ID NO 107; 394pp; English.  
XX  
CC The present invention relates to novel human coding sequences (ABQ99268-  
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in  
CC therapeutic, diagnostic and research methods. The polynucleotides may be  
CC used in the field of molecular biology as hybridisation probes, primers  
CC for PCR, for chromosome and gene mapping, for the recombinant production  
CC of protein, or in generation of anti-sense DNA or RNA. The  
CC polynucleotides are useful in diagnostics as expressed sequence tags  
CC (ESTs) for identifying expressed genes or for physical mapping of the  
CC human genome. The proteins may be used as molecular weight markers, or as  
CC nutritional sources or supplements. The proteins may be used to maintain  
CC and expand cell population in a totipotential or pluripotential state  
CC useful for re-engineering damaged or diseased tissues, transplantation,  
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
CC polynucleotides and proteins are useful for preventing, treating or  
CC ameliorating disorders involving aberrant protein expression or  
CC biological activity, e.g. haematopoietic disorders, central/peripheral  
CC nervous system diseases, mechanical and traumatic disorders, non-healing  
CC wounds, immune deficiencies and disorders, infectious diseases caused by  
CC viral, bacterial or fungal infection, autoimmune disorders, allergic  
CC reactions and conditions, coagulation disorders, or cancer. The  
CC polynucleotide sequences of the invention were assembled from ESTs  
CC isolated mainly by sequencing by hybridisation, and in some cases,  
CC sequences obtained from one or more public databases. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,21e-96 Length: 1413  
Score: 1206.50 Matches: 240  
Percent Similarity: 91.60% Conservative: 0  
Best Local Similarity: 91.60% Mismatches: 4  
Query Match: 89.17% Indels: 18  
DB: 6 Gaps: 1

US-09-939-853A-75 (1-261) x ABQ99374 (1-1413)

QY 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
DB 54 ATGGAGGTCTCCCGCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 113  
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
DB 114 CAAGCCAGGAGGACCTGTGACCATGGAGCAGAGAGAGCAAGGCCACAGCGCTGCGCTG 173

QY 41 GlySerPheProAlaGlyProAlaGluLeuSerLeuArgGluGlyProLeuThr 60  
 DB 174 GCGAGTTTCCCGGAGTGGCCCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 233  
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 234 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTCTGAACTCTCAGGCAGAGAGTAT 293  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluLysSer 100  
 DB 294 AACATCCCGAGCGTCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 353  
 QY 101 ArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuIle 120  
 DB 354 AGGGAGAAACAGAGAGAACTGTGTGTACCTGGGAACCTCTGGAGGGGCTTCTCATC 413  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 414 CCGGAGAGCCAGACAGGAGAGGCTCTTACTCTCTGTGAGTCCGCTCAGCCGCCCTGCA 473  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 DB 474 TCTCGGACCGGATCAGACACTACAGGATCCACTGCTTGACATGGCTGGCTGTATCATC 533  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 DB 534 TCACCGCGGCTCACCTTCCCTCACTCCAGGCGCTGGGGAGCCATTAC----- 581  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200  
 DB 582 -----TCTGAGGCTGGCCCGCTCCC 602  
 QY 200 cGlyLysAspIleProLeuProValThrValcInArgThrProLeuAsnTrpLysGluLe 220  
 DB 603 TGGCAAGGATATACCTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662  
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGlnSerLeuSerGluGln 240  
 DB 663 GGACAGCTCCCTCCCTGTTTCTGAAAGTGGCCAGGGAGGAGTCTTCTCATGTAGGG 722  
 QY 240 yLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAs 260  
 DB 723 TCTCGGAGTCCCTCAGCTTCTATCATCAGGCTGAATCAGGAGGCTGTCTTTGGATGA 782  
 QY 260 pAla 261  
 DB 783 TGCC 786  
 RESULT 7  
 ID AAL44090  
 XX AAL44090 standard; cDNA; 737 BP.  
 AC AAL44090;  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Mouse MARS short isoform protein coding sequence.  
 KW Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;  
 KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;  
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 XX  
 OS Mus sp.  
 XX  
 FH Location/Qualifiers  
 FT 1..633  
 FT /\*tag= a  
 FT /product= "Mouse MARS short isoform protein"  
 XX  
 PN WO200242452-A2.  
 XX

PD 30-MAY-2002.  
 XX 26-NOV-2001; 2001WO-CA001662.  
 XX 27-NOV-2000; 2000CA-02324663.  
 PR (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 PA  
 XX  
 XX  
 PI Mcglade JC, Loreto MP;  
 XX  
 XX WPI; 2002-566564/60.  
 DR P-PSDB; AAO15458.  
 XX  
 PT New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX  
 XX Claim 9; Page 77; 110pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present cDNA sequence encodes a mouse MARS protein.  
 XX  
 SQ Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 7,25e-96 Length: 737  
 Score: 1196.50 Matches: 240  
 Percent Similarity: 92.02% Conservative: 2  
 Best Local Similarity: 91.25% Mismatches: 2  
 Query Match: 88.43% Indels: 19  
 DB: 6 Gaps: 1  
 US-09-939-853A-75 (1-261) x AAL44090 (1-737)  
 QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerSerVal 20  
 DB 1 ATGGGAAGTCTGCCCCAGCAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 DB 61 CAAAGCCAGGACCTGTGACCATGGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120  
 QY 41 GlySerPheProAlaGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 DB 121 GGCAGTTTCCCGCAGGTGGCCCGCAGCTGTGCTGAGACTCGGGAGGCCATTGACC 180  
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 181 ATCGTCTCTGAGATGGAGACTGGTGGAGCGGTGCTGTCTGAAGTCTCAGGCAGAGAGAT 240  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluLysSer 100  
 DB 241 AACATCCCGAGGTCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 300  
 QY 101 ArgGluLysAlaGluLeuLeuLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuIle 120  
 DB 301 AGGGAGAAACAGAGAGAACTGTGTGTGTACCTGGGAACCTCTGGAGGGGCTTCTCATC 360  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 361 CCGGAGAGCCAGACAGGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCCGCCCTGCA 420  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 DB 421 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTTGTGACATGGCTGGCTGTATCATC 480

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 481 TCACCGCCCTCACCTCCCTCACTCCAGCCCTGGTGGACCATTC----- 528  
 QY 181 AspAspIleCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200  
 Db 529 -----TCTGAGGCTGGCCGCTCCC 549  
 QY 200 oGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220  
 Db 550 TGGCAAGGATATACCCCTACCTGACTGTGCGGAGGACACCACTCAACTGGAAGAGCT 609  
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGl 240  
 Db 610 GCACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTAGTGAGGG 669  
 QY 240 YLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA 260  
 Db 670 TCTCCGGAGTCCCTCAGCTTCTACATCAGCTGCAATGAGCGAGGCTGTCTCTTTGGATG 729  
 QY 260 spAla 261  
 Db 730 ATGCC 734  
 RESULT 8  
 ID AAS74750 standard; cDNA; 2049 BP.  
 AC AAS74750;  
 DT 13-FEB-2002 (first entry)  
 DE DNA encoding novel human diagnostic protein #10554.  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 OS Homo sapiens.  
 XX WO200175067-A2.  
 PN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US008631.  
 PF 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 XX WPI: 2001-639362/73.  
 DR P-PSDB; ABG10563.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 10554; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques for restoring normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences  
 XX  
 SQ Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;  
 Alignment Scores:  
 Pred. No.: 2,52e-89 Length: 2049  
 Score: 1128.50 Matches: 238  
 Percent Similarity: 89.51% Conservative: 1  
 Best Local Similarity: 89.14% Mismatches: 5  
 Query Match: 89.41% Indels: 23  
 DB: Gaps: 1  
 US-09-939-853A-75 (1-261) x AAS74750 (1-2049)  
 QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 Db 965 ATGGGAAGTCTCCCGAGGAGAGAAATCTTCCAGCCCAAGCTTGAGTTCTCTGTC 1024  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 1025 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAAAGCCACAGCGTGGCCCTG 1084  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLysGluProLeuThr 60  
 Db 1085 GGCAGTTTCCCGCAGGTGGCCCGGAGCTGCTCGAGACTCGGGAGGACCATGACC 1144  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 1145 ATCGTCTCTGAGATGGAGACTGTGTGGACGGTGTCTGAAGTCTCAGGCAGAGATAT 1204  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluLysSer 100  
 Db 1205 AACATCCCGCAGGCTCCAGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 1264  
 QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuIle 120  
 Db 1265 AGGGAAGACAGAGAACTGCTGTGTACTTGGAACTTGGAGGGGCTTCTCTCATC 1324  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 1325 CGGAGAGCCAGACCCAGGAGAGGCTCTTACTCTCTGTCTGCTCCGCTCAACGCCCTGCA 1384  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 Db 1385 TTCTGGACCCGATCAGACACTCAGGATTCATCGGCTTGACATGGCTGGCTGTACATC 1444  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 1445 TCACCGCGCTCACCTTCCCTCACTCCAGGCGCTGTGGACCATTC----- 1492  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200  
 Db 1493 -----TCTGAGGCTGGCCGCTCCC 1513  
 QY 200 oGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220  
 Db 1514 TGGCAAGGATATACCCCTACCTGACTGTGCGGAGGACACCACTCAACTGGAAGAGCT 1573  
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGluLysSerLeuLeuSerGlu 239  
 Db 1574 GGACAGTCCCTCTCTTCTGAAGCTGCCACAGGGAGGAGTCTCTTCTCAGAGGAG 1633  
 QY 240 GlyLeuArgGluSer-LeuSerPheTyrIleSer-LeuAsnAspGluAlaValSer-Leu 258

Db 1634 GGGCTCGGAGTCCCTCAGCTTCTACATCAGCCCTGATGACGAGGCTGTCTCTTTTG 1693

Qy 259 AspA261  
 ID AAL44087  
 Db 1694 GATGATGCC 1702

RESULT 9  
 AAL44087  
 ID AAL44087 standard; cDNA; 1348 BP.  
 XX AAL44087;  
 AC AAL44087;  
 XX  
 DT 03-OCT-2002 (first entry)  
 XX  
 DE Mouse modulator of antigen receptor signalling protein coding sequence.  
 XX  
 KW Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;  
 KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;  
 KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;  
 KW immunosuppression; myeloproliferative disorder; breast cancer.  
 XX  
 OS Mus sp.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 282..1061  
 FT /tag= a  
 FT /product= "Mouse MARS protein"  
 XX  
 XX WO200242452-A2.  
 XX  
 XX 30-MAY-2002.  
 XX  
 XX 26-NOV-2001; 2001WO-CA001662.  
 XX  
 XX 27-NOV-2000; 2000CA-02324663.  
 XX  
 XX (HOSP-) HOSPITAL FOR SICK CHILDREN.  
 XX  
 XX Mcglade JC, Loreto MP;  
 XX  
 XX WPI; 2002-566564/60.  
 DR P-PSDB; AA015456.  
 XX  
 XX New isolated modulator of antigen receptor signaling protein or its  
 PT fragment, useful for treating malignant disorders such as myeloid  
 PT malignancies, autoimmune disorders and myeloproliferative disorders.  
 XX  
 XX Claim 10; Fig 1A; 110pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of modulator  
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a  
 CC putative tumour suppressor gene and exhibits structural and sequence  
 CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and  
 CC protein sequences of the invention are useful for the treatment of  
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune  
 CC disorders, immunosuppression, myeloproliferative disorders and  
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.  
 CC breast cancer). The present cDNA sequence encodes a mouse MARS protein  
 XX  
 XX Sequence 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 9,84e-81 Length: 1348  
 Score: 1028.00 Matches: 208  
 Percent Similarity: 85.50% Conservative: 16  
 Best Local Similarity: 79.39% Mismatches: 34  
 Query Match: 75.98% Indels: 4  
 DB: 6 Gaps: 3

US-09-939-853a-75 (1-261) x AAL44087 (1-1348)

Qy 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 ID AAL44087  
 DB: 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;

Db 282 ATGGGAGTTTGTCCAGCAGAGGGAAAC---TCCAGCCCCAGCCCCAGCTCTCTGTGT 338

Qy 21 GlnGlyInGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 339 CCAGACCAGGAACCCGTGTCCATGCAACACAGAAAGACACAAGGTCACAGCTGTGGCCCTG 398

Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 399 GGCAGTTTCCACAGAGTGAACAGGCCAGACTATCTCTGAGACTCGGGAGCCGCTGACC 458

Qy 61 IleValSerGluAspGlyAspThrIleThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 459 ATCATCTCTGAGGATGAGATTGGTGGACAGTCCAGTCGGAAGTCTCAGGCAGAGGTAC 518

Qy 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 Db 519 CATATGCCCGAGTGTATGGTAAAGTCGCCACGGGTGGCTGTACGAGGCCCTGAGC 578

Qy 101 ArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 Db 579 CGGAGAAAGCCGAGGAACCTACTCTGTGTACCTGGGAACCCCGAGGGGCTTCTCATC 638

Qy 121 ArgGluSerClnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 639 CGGAGAGCCAGACCCAGGAGGCTGCTATTCCTGTCCGTCCGACTCAGCCGCCCTGCA 698

Qy 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160  
 Db 699 TCTTGGACCGGATCAGACACTACAGGATACAGCGCTTGACAATGGCTGGCTGTACATC 758

Qy 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 Db 759 TCACCTCGCCTCACCTTCCCTCCTCCTCCAGCGCTTGTGGAGCATTACTCTGAGTAGCA 818

Qy 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 Db 819 GATGGCATCTGTGTCCCTCAGGAGCGGTGTCTCTGAGAGAGCTTGGCCACTACCT 878

Qy 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 Db 879 GGCAAGAGATACACCTCCACCTGTGACTGTGCCAACATCATCAATAATGGAAAGCTG 938

Qy 221 AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuLeuSerGlu 239  
 Db 939 GACCGAGCCTCTCTGTCTCGGAGACACCTCGAGTGGGAGGCATCTCTGCTCAGTGAG 998

Qy 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259  
 Db 999 GGGCTCCGAGAGTCCCTCAGTTCCTATCATCAGCCTGGCTGAGGAC-----CCCTTGGAT 1052

Qy 260 AspAla 261  
 Db 1053 GATGCT 1058

RESULT 10  
 ABQ98670  
 ID ABQ98670 standard; DNA; 763 BP.  
 XX  
 XX ABQ98670;  
 AC  
 XX  
 DT 04-NOV-2002 (first entry)  
 XX  
 DE Human ORF477 coding sequence.  
 XX  
 KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antiinflammatory; gene therapy; human; ORFX; atherosclerotic plaque;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2002082206-A1.  
 PN

XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX PR 30-MAY-2000; 2000US-0208427P.  
 XX PA (LEAC/) LEACH M D.  
 XX PA (MEHR/) MEHRABAN F.  
 XX PA (CONLEY/) CONLEY P B.  
 XX PA (TOPP/) TOPPER J N.  
 XX PA (LAWD/) LAW D.  
 XX PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX WPI; 2002-626554/67.  
 XX DR P-PSDB; ABP64107.  
 XX PT New polypeptide designated ORFX are present in human atherogenic cells  
 XX PT and are useful to prevent and treat ORFX-associated disorders including  
 XX PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 XX PT inflammatory disease.  
 XX PS Claim 2; SEQ ID NO 953; 78pp; English.  
 XX CC The present invention relates to novel human ORFX polypeptides and their  
 XX CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 XX CC were discovered in human atherogenic cells, in particular in platelets  
 XX CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 XX CC many other tissues as well. Atherogenic cells are cells which have the  
 XX CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 XX CC nucleic acids are useful for treating or preventing a pathological  
 XX CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 XX CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 XX CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 XX CC for this patent did not form part of the printed specification, but was  
 XX CC obtained in electronic format directly from the USPTO web site at  
 XX CC seqdata.uspto.gov/sequence.html?docID=20020082206  
 XX SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,59e-63 Length: 763  
 Score: 826.00 Matches: 158  
 Percent Similarity: 99.37% Conservative: 0  
 Best Local Similarity: 99.37% Mismatches: 1  
 Query Match: 61.05% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-939-853A-75 (1-261) x ABQ98670 (1-763)  
 QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
 DB 286 ATGGGAAGTCTGCCAGAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 345  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 DB 346 CAAAGCCAGGACCTGTACCATGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 405  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 DB 406 GGCAGTTTCCCGCAGGTGGCCCGCCGAGTGTCTGAGACTCGGGGAGCCATTGACC 465  
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 466 ATCGTCTCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 525  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 DB 526 AACATCCCGAGGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 585  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120

DB 586 AGGAGAGAAAGCAGAGAGAACTGCTGTTGTACCTGGGAAACCTGGAGGGGCTTCTCATC 645  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 646 CGGAGAGCCAGAGCCAGGAGAGGCTTCTTACTCTGTCTGTCAGTCCGCTCAGCGCCCTGCA 705  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyr 159  
 DB 706 TCTTGGGACCGGATCAGACACTACAGGATCCACTGCTTGACATGGCTGGCTGTAC 762  
 RESULT 11  
 AAS74748  
 ID AAS74748 standard; cDNA; 603 BP.  
 XX AC AAS74748;  
 XX DT 13-FEB-2002 (first entry)  
 XX DE DNA encoding novel human diagnostic protein #10552.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 XX DR P-PSDB; ABG10561.  
 XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 XX PT diagnostics, forensics, gene mapping, identification of mutations  
 XX PT responsible for genetic disorders or other traits and to assess  
 XX PT biodiversity.  
 XX PS Claim 1; SEQ ID NO 10552; 103pp; English.  
 XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 XX CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 XX CC and in recombinant production of (II). The polynucleotides are also used  
 XX CC in diagnostics as expressed sequence tags for identifying expressed  
 XX CC genes. (I) is useful in gene therapy techniques to restore normal  
 XX CC activity of (II) or to treat disease states involving (II). (II) is  
 XX CC useful for generating antibodies against it, detecting or quantitating a  
 XX CC polypeptide in tissue, as molecular weight markers and as a food  
 XX CC supplement. (II) and its binding partners are useful in medical imaging  
 XX CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 XX CC involving aberrant protein expression or biological activity. The  
 XX CC polypeptide and polynucleotide sequences have applications in  
 XX CC diagnostics, forensics, gene mapping, identification of mutations  
 XX CC responsible for genetic disorders or other traits to assess biodiversity  
 XX CC and to produce other types of data and products dependent on DNA and  
 XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 XX CC coding sequences of the invention. Note: The sequence data for this  
 XX CC patent did not appear in the printed specification, but was obtained in  
 XX CC electronic format directly from WIPO at  
 XX CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;  
 Alignment Scores:

Pred. No.: 4,2e-55 Length: 603  
 Score: 731.00 Matches: 151  
 Percent Similarity: 82.45% Conservative: 4  
 Best Local Similarity: 80.32% Mismatches: 21  
 Query Match: 54.03% Indels: 12  
 DB: 5 Gaps: 4

US-09-939-853A-75 (1-261) x AAS74748 (1-603)

QY 78 ArgGluTyrAsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGlu 97  
 Db 61 CGCAGGGGAACTCCCAAGAGTCGGTG---AAGATCCTCCGGGACTGGCTGTAC--- 114  
 QY 98 GlyLeuSerArgGluLysAlaA-----GluGluLeuLeuLeuProGlyAsn 113  
 Db 115 ---TTGCACCCCTACAAGCCTTACCCTCAGAGGAGGAGCTGAGCCCTTCTGGACAG 171  
 QY 114 ProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTyrSerLeuSer 133  
 Db 172 ACCAACCTGTGCAGTGCTG-----CAAGACATGGCTCTTACTCTCTGTCA 216  
 QY 134 ValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIleHisCysLeu 153  
 Db 217 GTCCSCCTCAGCCGCCCTCGATCTCTGGGACCGGATCAGACACTACAGATCCACTGGCCTT 276  
 QY 154 AspAsnGlyTyrLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuVal 173  
 Db 277 GACATGCTGGCTGTATCATCTCACCGCGCTCACCTTCCCTCACTCCAGGCCCTGGTG 336  
 QY 174 AspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysValLeu 193  
 Db 337 GACCAATTACTGTAGCTGGCGGATGACATCTGCTGCTTACTCAAGGAGCCCTGTGTCTGT 396  
 QY 194 GlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThr 213  
 Db 397 CAGAGGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACCTGTGACTGTGCAGAGACA 456  
 QY 214 ProLeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGlu 233  
 Db 457 CCACCTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAG 516  
 QY 234 GluSerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp 253  
 Db 517 GAGTCTCTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTATCATCAGCTGATGAC 576

254 GluAlaValSerLeuAspAla 261  
 577 GAGGCTGTCTCTTGGATGATGCC 600

RESULT 12  
 ABQ99151  
 ID ABQ99151 standard; DNA; 875 BP.  
 XX AC ABQ99151;  
 XX DT 04-NOV-2002 (first entry)  
 XX DE Human ORF958 coding sequence.  
 XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;  
 KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.  
 XX OS Homo sapiens.  
 XX PN US2002082206-A1.  
 XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX ID ABX62975 standard; cDNA; 3756 BP.

PR 30-MAY-2000; 2000US-0208427P.  
 XX (LEAC/) LEACH M D.  
 PA (MEHR/) MEHRABAN F.  
 PA (CONL/) CONLEY P B.  
 PA (TOPP/) TOPPER J N.  
 PA (LAWD/) LAW D.  
 XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 PI WPI; 2002-626554/67.  
 XX P-PSDB; ABP64588.  
 DR New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 XX Claim 2; SEQ ID NO 1915; 78pp; English.  
 XX The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?DocID=20020082206  
 XX SQ Sequence 875 BP; 205 A; 259 C; 225 G; 185 T; 0 U; 1 Other;  
 Alignment Scores:  
 Pred. No.: 3,66e-42 Length: 875  
 Score: 586.00 Matches: 112  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 43.31% Indels: 0  
 DB: 6 Gaps: 0

US-09-939-853A-75 (1-261) x ABQ99151 (1-875)

QY 150 IleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeuThrPheProSerLeu 169  
 Db 4 ATCCACTGGCTTGACAAATGGCTGGCTGTATCATCTCACCGCGCTCACCTTCCCTCACTC 63  
 QY 170 GlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLysGlu 189  
 Db 64 CAGGCCCTGGTGACCATTTACTTCTGAGCTGGCGGATGATCTGCTGCTATCTCAGGAG 123  
 QY 190 ProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThr 209  
 Db 124 CCCTGTGTCTCTGACAGAGGCTGGCCCGCTCCCTGGCAAGATATACCCCTACCTGTGACT 183  
 QY 210 ValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAla 229  
 Db 184 GTGCAGAGACACCACTCAACTGGAAGAGCTGGACAGCTCCCTCTGTTTCTTGAGCT 243  
 QY 230 AlaThrGlyGluGluSerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIle 249  
 Db 244 GCCACAGGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTACATC 303  
 QY 250 SerLeuAsnAspGluAlaValSerLeuAspAla 261  
 Db 304 AGCTGATGATGACGAGGCTGTCTCTTTGGATGATGCC 339

RESULT 13  
 ABX62975  
 ID ABX62975 standard; cDNA; 3756 BP.

XX AC ABX62975;  
 XX DT 25-FEB-2003 (first entry)  
 XX DE Human activated T cell cDNA #91.  
 XX KW T cell; gene; ss; differential expression; T cell activation;  
 KW antiallergic; cytostatic; immunosuppressive; antimicrobial; gene therapy;  
 KW allergy; cancer; graft versus host disease; infection;  
 KW autoimmune disorder.  
 XX OS Homo sapiens.  
 XX PN US2002137077-A1.  
 XX PD 26-SEP-2002.  
 XX PF 25-OCT-2001; 2001US-00002600.  
 XX PR 25-OCT-2000; 2000US-0243521P.  
 XX PA (HOPK/) HOPKINS C M.  
 XX PA (PETE/) PETERSON D P.  
 XX PA (COCK/) COCKS B G.  
 XX PA (HAWK/) HAWKINS P R.  
 XX PI Hopkins CM, Peterson DP, Cocks BG, Hawkins PR;  
 XX DR WPI; 2003-102381/09.  
 XX PT New combination comprising several cDNAs that are differentially  
 PT expressed in activated T cells, useful for diagnosing, treating, staging  
 PT or monitoring treatment for allergy, cancer, infectious and/or autoimmune  
 PT disorders.  
 XX PS Claim 1; Page: 180pp; English.  
 XX CC This invention relates to the sequences of several cDNAs that are  
 CC differentially expressed in activated T cells. The sequences of the  
 CC invention may have antiallergic, cytostatic, immunosuppressive and  
 CC antimicrobial activity and may be used in gene therapy. The invention  
 CC also comprises a method for screening samples for differentially  
 CC expressed genes and a method for detecting these cDNAs by hybridisation.  
 CC The methods and compositions of the present invention are useful for  
 CC diagnosing, treating, staging or monitoring treatment for allergy,  
 CC cancer, chronic graft versus host disease, infectious and/or autoimmune  
 CC disorders. The present sequence represents a cDNA of the invention that  
 CC is differentially expressed in activated T cells  
 XX SQ Sequence 3756 BP; 1007 A; 918 C; 960 G; 871 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9.58e-33 Length: 3756  
 Score: 488.00 Matches: 101  
 Percent Similarity: 57.20% Conservative: 46  
 Best Local Similarity: 39.30% Mismatches: 94  
 Query Match: 36.07% Indels: 16  
 Gaps: 3

US-09-939-853A-75 (1-261) x ABX62975 (1-3756)

QY 5 ProSerArgArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24  
 DB 1098 CCAGGAAAGAAAGAAATGGGAAACAGCATGAATCCACCCCTGCGCTGCGGAGG 1157  
 QY 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44  
 DB 1158 CCCCTGCCCAACCCGAGGAGGACTGGATAGCGACTTCCTTCCGCTGCTAAGTACTACCCG 1217  
 QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgGluGlyProLeuThrIleValSerGlu 64  
 DB 1218 TCTCTGCATCAGCCCCCGATATTCGCGGAGGAGAAATGCTGTGATTTCTGAT 1277

QY 65 AspGlyAspTrrTrrValLeuSerGluValSerGlyArgGluTyrAsnIleProSer 84  
 DB 1278 GAAGGGGCTGGTGGAAAGCTATTCTTAGCACTGGTGGAGAGAGTACATCCCTGGA 1337  
 QY 85 ValHisValGlyLysValSerHisGlyTrrLeuTyrGluGlyLeuSerArgGluValAla 104  
 DB 1338 ATATGTGGCCAGAGTTTACATGGCTGGCTGTTTGGAGGCTGGGAGAGAGAGAGCC 1397  
 QY 105 GluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuIleArgGluSerGln 124  
 DB 1398 GAGGAGCTGCTGCAGTCCAGACACAAAGCTGGCTCTTCATGATCAGAGAGAGTGA 1457  
 QY 125 ThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144  
 DB 1458 ACCAAGAAAGGTTTACTCTGCTGGGTGAGACAAAGG-----CAG 1499  
 QY 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTrrLeuTyrIleSerProArgLeu 164  
 DB 1500 GTAAAGCATTTACCGCATTTTCGCTGCTGCCCAACACTGGTACTACATTTCCCGAGGCTC 1559  
 QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspIleCys 184  
 DB 1560 ACCTTCCAGTCCGTCGAGGACCTGGTGAACCACTATTCTGAGTGGCTGTGATGGCTGTC 1619  
 QY 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
 DB 1620 TGTGTGCTCACCAGCCCTGCTGACACAAAGACGCGCTGCCAGCAGTGCAGGCTCC 1679  
 QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTrrLysGluLeuAspSerSerLeu 224  
 DB 1680 AGCTCACCTGCACCTTCGTCAGAAAGACTGTGAGTGGAGAGAGTGTCCAGA----- 1733  
 QY 225 LeuPheSerGluAlaAlaThrGly-----GluGluSerLeuLeu 237  
 DB 1734 ---CTGCAGGAGGACCCCGAGGACAGAGAACCCGCTTGGGTGAGAGCTCCCTTTTC 1790  
 QY 238 SerGluGlyLeuArgGluSerLeuSerPheTrrIleSerLeuAsnAspGlu 254  
 DB 1791 AGCTATGGCTTCGAGAGAGCATTGCCTCTTACCTGTCCCTGACCCAGTGA 1841

RESULT 14  
 AAS02049  
 ID AAS02049 standard; cDNA; 2109 BP.  
 AC AAS02049;  
 XX 16-JUL-2001 (first entry)  
 DT DNA encoding molecule for disease detection and treatment, mddt14.  
 DE Human; mddt14; gene therapy; adenosine deaminase deficiency; ADA;  
 KW severe combined immunodeficiency syndrome; cystic fibrosis; thalassemia;  
 KW familial hypercholesterolemia; haemophilia; factor VIII; factor IX;  
 KW cancer; cell proliferation; parasite; human retrovirus; HIV; hepatitis B;  
 KW hepatitis C; Candida albicans; Plasmodium falciparum;  
 KW Paracoccidioides brasiliensis; Trypanosoma brasiliensis; ss.  
 OS Homo sapiens.  
 XX WO200123538-A2.  
 PN 05-APR-2001.  
 PD 22-SEP-2000; 2000WO-US026085.  
 PF 28-SEP-1999; 99US-0156565P.  
 PR 30-NOV-1999; 99US-0168197P.  
 XX (INCYTE) INCYTE GENOMICS INC.  
 XX Hodgson DM, Lincoln SE, Russo PD, Spiro PA, Banville SC;  
 PI Bratcher SR, Dufour GE, Cohen HZ, Rosen BH, Shah P, Chalup MS;



PR 02-OCT-2000; 2000US-0237295P.  
PR 02-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
XX (AVAL-) AVALON PHARM.  
XX  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX  
DR WPI; 2002-188264/24.  
XX  
XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX  
PS Claim 1; SEQ ID NO 3526; 44pp; English.  
XX  
XX The present invention describes a method (M1) for screening for an anti-  
CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms's  
CC tumour  
XX  
SQ Sequence 2665 BP; 736 A; 617 C; 689 G; 623 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 7.55e-33 Length: 2665  
Score: 487.00 Matches: 101  
Percent Similarity: 57.20% Conservative: 46  
Best Local Similarity: 39.30% Mismatches: 94  
Query Match: 35.93% Indels: 16  
DB: 6 Gaps: 3

US-09-939-853A-75 (1-261) x ABL65189 (1-2665)

QY 5 ProSerArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGln 24  
Db 24 CCAGGAAAGAAAGAAATGGGAACAGCATGAATCCACCCCTCGCGCTCCGAGAGG 83  
QY 25 ProValThrMetGluAlaGluArgSerLysLysLysLysLysLysLysLysLysLys 44  
Db 84 CCCCTGCCCAACCCCGAGGAGCTGGATAGCGACTTCCTTGGCTGCTAAGTACTACCCG 143  
QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLysGluProLeuThrIleValSerGlu 64  
Db 144 TCTCTGACATCAGCCCCCGATATCCCGAGGAGAACTGCGTGTGATTCTGAT 203  
QY 65 AspGlyAspTyrThrValLeuSerGluValSerGluValSerGluValSerGluValSer 84  
Db 204 GAAGGGGGCTGGTGAAGACTATTTCTTCTTACACTGCTCGAGAGAGTTACATCCCTGGA 263  
QY 85 ValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSerArgGluLysAla 104  
Db 264 ATATGCTGGCCAGAGTTTACCATGGCTGGCTGTTTGGGGCTGGGCGAGAGCAAGGCC 323

QY 105 GluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGln 124  
Db 324 GAGGAGCTGCTCAGCTGCCAGACACAAAGGTGGCTCCTTCATCATCAGAGAGAGTGAG 383  
QY 125 ThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTyrAspArg 144  
Db 384 ACCAAGAAGGGTTTACTACTCTGCGTGAGA-----CACAGGCGAG 425  
QY 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIleSerProArgLeu 164  
Db 426 GTAAAGCATTACCGCATTTTCGCTGCGCAACAACTGGTACTACATTTCCCGAGGCTC 485  
QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCys 184  
Db 486 ACCTTCCAGTGGCTGGAGACCTGGTGAACCACTATTCTGAGTGGCTGATGGCTGTGC 545  
QY 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
Db 546 TGTGTGCTCACCACGCCCTGCTGACACAAAGCAGCGCTGCCACAGCAGTGGGGCTCC 605  
QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLeuAspSerSerLeu 224  
Db 606 AGCTCACCTGTTCACCTTGGCTCAGAAAGACTGTGGACTGGAGAGAGTGTCCAGA----- 659  
QY 225 LeuPheSerGluAlaAlaThrGly-----GluGluSerLeuLeu 237  
Db 660 ---CTGAGGAGGACCCCGAGGAAACAGAGAACCCGCTGGGGTAGACGAGTCCCTTTTC 716  
QY 238 SerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu 254  
Db 717 AGCTATGGCTTCGAGAGAGCAATTCCTTACCTGTCCCTGACCGAGTGAG 767

Search completed: February 20, 2004, 09:56:47  
Job time : 377 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:39:38 ; Search time 2518 Seconds  
(without alignments)  
3095.325 Million cell updates/sec

Title: US-09-939-853A-75  
Perfect score: 1353  
Sequence: 1 MGSIPSRKSLPSPSLSSV.....RESLSFVLSLNDVSLDDA 261

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/cn2.1/USPTO.spool.p/US09939853/runat.19022004.145339.24496/app.query.fasta.1.455  
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human4.cd1 -LIST=45  
-DOCALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09939853 @CGN 1 1 4237 @runat.19022004.145339.24496 -NCPU=6 -ICPU=3  
-NO MMAP -LARGSEQ=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpi:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1273	94.1	1002	12	BQ052308	BQ052308 AGENCOURT
2	1183	87.4	1069	12	BQ052468	BQ052468 AGENCOURT
3	1148.5	84.9	1020	12	BQ054281	BQ054281 AGENCOURT
4	1028	76.0	2837	11	AK088672	AK088672 Mus muscu
5	1024	75.7	2974	11	AK030877	AK030877 Mus muscu
6	933.5	69.0	926	11	AK020837	AK020837 Mus muscu
7	878.5	64.9	660	13	BY742155	BY742155 Mus muscu
8	867	64.1	1201	9	AL541041	AL541041 AL541041
9	711	52.5	566	12	BQ284179	BQ284179 602408226
10	680.5	50.3	986	12	BQ054265	BQ054265 AGENCOURT
11	611.5	45.2	960	13	BQ944126	BQ944126 AGENCOURT
12	560.5	41.4	878	12	BQ053486	BQ053486 AGENCOURT
13	557	41.2	377	9	AA959151	AA959151 v251906.r
14	556.5	41.1	660	10	BB635615	BB635615 BB635615
15	555	41.0	597	9	AL844311	AL844311 AL844311
16	546	40.4	778	12	BG178487	BG178487 602328305
17	545	40.3	606	9	AL844309	AL844309 AL844309
18	543	40.1	614	9	AL844307	AL844307 AL844307
19	531	39.2	569	14	CB426333	CB426333 601508 MA
20	517.5	38.2	627	10	BB619854	BB619854 BB619854
21	492	36.4	2810	11	AK036167	AK036167 Mus muscu
22	491.5	36.3	831	29	AY416279	AY416279 Mus muscu
23	491	36.3	1386	11	AK041565	AK041565 Mus muscu
24	489.5	35.2	831	29	AY416277	AY416277 Homo sapi
25	478	35.3	1201	9	AL539427	AL539427 AL539427
26	477.5	35.3	655	14	CF913437	CF913437 A0648D04-
27	476.5	35.2	1997	11	AK037901	AK037901 Mus muscu
28	473	35.0	1201	9	AL551370	AL551370 AL551370
29	470.5	34.8	701	13	BX849096	BX849096 BX849096
30	468	34.6	1106	13	BX436423	BX436423 BX436423
31	467	34.5	972	13	BQ707614	BQ707614 AGENCOURT
32	464	34.3	1133	9	AL549826	AL549826 AL549826
33	452.5	33.4	871	13	BQ436143	BQ436143 AGENCOURT
34	447.5	33.1	775	14	CB938697	CB938697 IPCGJX13
35	442	32.7	1063	13	BQ072745	BQ072745 AGENCOURT
36	441	32.6	723	12	BM950089	BM950089 UI-M-BHOP
37	439.5	32.5	777	14	CB938723	CB938723 IPCGJX13
38	439.5	32.5	1003	13	BX415149	BX415149 BX415149
39	434	32.1	792	14	CA366437	CA366437 641999 NC
40	433	32.0	767	14	CF287401	CF287401 AGENCOURT
41	433	32.0	794	12	BQ677567	BQ677567 602624118
42	433	32.0	849	12	BI769183	BI769183 603053793
43	425.5	31.4	603	13	BQ553005	BQ553005 H4019E02-
44	424	31.3	570	9	AI471720	AI471720 tal6a01.x
45	422	31.2	673	10	BB638252	BB638252 BB638252

# ALIGNMENTS

RESULT 1  
BQ052308  
LOCUS BQ052308  
DEFINITION AGENCOURT\_6868571 NIH\_MGC\_106 Homo sapiens CDNA clone IMAGE:5933542  
ACCESSION BQ052308  
VERSION BQ052308.1 GI:19811648  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1002)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>;  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2118 row: d column: 23  
 High quality sequence stop: 670.

FEATURES  
 source  
 1..1002  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5933542"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 106"  
 /note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,246-118 Length: 1002  
 Score: 1273.00 Matches: 256  
 Percent Similarity: 97.36% Conservative: 2  
 Best Local Similarity: 96.60% Mismatches: 3  
 Query Match: 94.09% Indels: 4  
 DB: 12 Gaps: 0

US-09-939-853A-75 (1-261) x BQ052308 (1-1002)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 DB 96 ATGGGAAGTCTCCAGAGAGAAATCTTCCACGCCCAAGCTTGAAGTCTCTGTC 155  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAla 40  
 DB 156 CAAGGCCAGGACCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTG 215  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyProLeuThr 60  
 DB 216 GGCAGTTTCCCGCAGGTGCCCGCCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 275  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 DB 276 ATCGTCTCTGAGATGAGACTGGTGAGCGTGTCTGAAGTCTCAGGCAGAGATAT 335  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 DB 336 AACATCCCGCAGGTCCACGTGCGCAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC 395  
 QY 101 ArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyValAlaPheLeuIle 120  
 DB 396 AGGAGAAAGCAGAGAACTGTGTGTGTATCTGGACCCCTGGAGGGCCCTTCTCATC 455  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 DB 456 CCGGAGAGCAGACACAGGAGAGCTTACTTCTGTCTGTCAGTCCGCTCAGCCGCGCTCA 515  
 QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAsnGlyTrpLeuTyrIle 160

DB 516 TCCTGGACCGGATCAGACACTACAGGATCTACTGCTTGGACATGGCTGGCTGTACATC 575  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180  
 DB 576 TCACCGCGCTCACCTTCCTCTCACTCAGGCGCTGGTGGACCATTAATCTTGAAGCTGGC 635  
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200  
 DB 636 GATGACATCTGTGCTACTCAAGAGCGCTGTGCTTGCAGAGGCTGGCCCGTCCCT 695  
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220  
 DB 696 GGCRAAGGATATACCTCTACTGTGACAGGACACCACTCACTGGAAGAGCTG 755  
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGluGluSerLeuLeuSerGlu-G 240  
 DB 756 GACAGCTCCCTCTCTGTTTCTGAAACTGCCACAGGGGAGGAGTCTCTTCTCAGTGAGG 815  
 QY 240 LysLeuArg-GluSerLeuSerPheTyrIleSerLeuAsnAspGlu-AlaValSerLeuAs 259  
 DB 816 GTCTCCGGGAGTCCCTCAGCTTCTACATCAGCCCTGATGACGAGGCTGTCTCTTGA 875  
 QY 259 pAspAla 261  
 DB 876 TGAAGCC 882

RESULT 2  
 BQ052468  
 LOCUS AGENCOURT 6868422 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5933772  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BQ052468  
 VERSION BQ052468.1 GI:19811808  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 1069)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2118 row: n column: 13  
 High quality sequence stop: 681.  
 Location/Qualifiers  
 1..1069  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5933772"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 106"  
 /note="Organ: blood; Vector: pOTB7; Site: 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

FEATURES  
 source  
 1..1069  
 Location/Qualifiers

## ORIGIN

## Alignment Scores:

```

Pred. No.: 3,47e-109 Length: 1069
Score: 1193.00 Matches: 243
Percent Similarity: 92.05% Conservative: 0
Best Local Similarity: 92.05% Mismatches: 2
Query Match: 87.44% Indels: 20
DB: 12 Gaps: 1

US-09-939-853A-75 (1-261) x BQ052468 (1-1069)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 89 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCCTCTGTC 148
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluAuctSerLysAlaThrAlaValAlaLeu 40
Db 149 CAGGGCAGGAGCTGTGACCATGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 208
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 209 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTCTCGACTCGGGAGGACATTGACC 268
QY 61 IleValSerClnuAspGlyAspTrrThrValLeuSerClnuValSerGlyArgGluTyr 80
Db 269 ATCGTCTCTGAGGATGGAGACTGGTGAGCGTGTCTCTGAACTCTCAGGCGAGAGATAT 328
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrrLeuTyrGluGlyLeuSer 100
Db 329 AACATCCCGAGCTCCAGCTGGCCAAAGTCTCCATGGTGTATGAGGGCCCTGAGC 388
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeulle 120
Db 389 AGGAGAAACAGAGAACTGCTGTGTTTACCTGGGAACCTCGAGGGCCCTTCCTCATC 448
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 449 CGSAGAGCCAGACCCAGGAGAGGCTTTACTCTCTCTGAGTCCGCTCCAGCCGCTGCA 508
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrrLeuTyrIle 160
Db 509 TCCTGGGACCGGATCAGACACTACAGGATCCACTGCTGACAAATGGCTGTATCATC 568
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeu-ValAspHisTyrSerGluLeuAl 180
Db 569 TCACCGGGCCCTACCTTCCCTCCTCCAGCCGCTGGTGGACCATTTACTCTGAGGG--- 625
QY 180 aAspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPr 200
Db 626 -----GCTGGCCCGCTCCC 639
QY 200 oGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220
Db 640 TGGCAAGGATATACCCCTACTCTGACTGTGCAGAGGACACCACTCACTGGAAAGAGCT 699
QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGluGluSerLeuLeuSerGluG 240
Db 700 GGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGGAGGAGTCTCTCTCAGTGAGG 759
QY 240 lyLeuArgGlu-SerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 760 GTCTCCGGGAAGTCCCTCAGCTTCTACATCAGCCTCAATGACGAGGCTGTCTTTTGGAT 819
QY 260 AspAla 261
Db 820 GATGCC 825

RESULT 3
BQ054281
LOCUS
DEFINITION AGENCOURT_6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
5', mRNA sequence.
ACCESSION BQ054281
VERSION BQ054281.1 GI:19813621
KEYWORDS EST.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 1020) Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mhc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/MCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2125 row: j column: 11
High quality sequence stop: 556.
FEATURES
Location/Qualifiers
1..1020
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 1,02e-105 Length: 1020
Score: 1148.50 Matches: 237
Percent Similarity: 93.13% Conservative: 7
Best Local Similarity: 90.46% Mismatches: 11
Query Match: 84.89% Indels: 7
DB: 1 Gaps: 1
US-09-939-853A-75 (1-261) x BQ054281 (1-1020)
QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 212 ATGGGAAGTCTGCCAGCAGAGAAAATCTCTGCCAAGCCCAAGCTTGAGTTCCCTCTGTC 271
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 272 CAAGCCAGGAGACCTGTGACCATGCAAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 332 GGCAGTTTCCCGCAGGTGGCCCGCGGAGCTGTCTGAGACTCGGGAGGCCATTGACC 391
QY 61 IleValSerGluAspGlyAspTrrThrValLeuSerGluValSerGlyArgGluTyr 80
Db 392 ATCGTCTCTGAGGATGGAGACTGGTGGACGCTGTCTGAACTCTCAGGCGAGAGATAT 451
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrrLeuTyrGluGlyLeuSer 100
Db 452 AACATCCCGAGCTCCAGCTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGGCCCTGAGC 511
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeulle 120
Db 512 AGGAGAAAGCAGAGGAACTGCTGTGTTTACCTNGAACCTCGAGGGGCTTCTCTCATC 571
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140

```

```

Db      572 CGGAGAGCCAGACAGAGAGGCTCTTACTCTCTGTCAGTCCGCCCTCAGCCGCCCTGCA 631
Qy      141 SerTrpAspArgIleAeGHisTyArgIleHisCysLeuAspAsnGlyTTPLeuTyrlle 160
Db      632 TCCTGNGACCGGATCAGACACTACAGATCCAGCTGCCTTGACAATGGCTGGCTGTATCATC 691
Qy      161 SerProArgLeuThrPheProSerLeuGluAla-LeuValAspHisTyrSerGluLeuAl 180
Db      692 TCACGGCGCCTCACCTTCCCTCTACTCCAGGCCCTGTGTGACCATTTACTTCTGAGCTGGC 751
Qy      180 aAspAspileCysCysLeu-LeuLysGluProCysValLeuGluAlaGlyProLeuLeuP 200
Db      752 GGATGACATCTGCTGCTTANCTCAGGACCCCTGTCTGCTGCAAGGGCTGCCCGCTCC 811
Qy      200 roGlyLysAspile-ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGlu 219
Db      812 CTGGCAAGATATACCCCTACCTGTGACTGTGCAAGGACACCATCTACTCGGAGAG 871
Qy      220 LeuAspSerSerLeuLeu---PheSerGluAlaAlaThrGlyGluGluSerLeuLeuSer 238
Db      872 CTGGACAGCTCCCTCTGTTTCTTGAAGCTGCCCCAGGGGAGGAGTCTCTTCATCATT 931
Qy      239 Gludly--LeuArgGluSerLeuSerPheTyrlleSerLeu-AenAspGluAla 255
Db      932 GAGGGGTCTCCGGAGTCCCTCAGCTTCTACATCAACCTGTATGACACAGGCT 985

RESULT 4
LOCUS   AK088672                2637 bp     mRNA      linear      HTC 20-SEP-2003
DEFINITION Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430023D24 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.
ACCESSION AK088672
VERSION   AK088672.1 GI:26353729
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS   Carninci,P. and Hayashizaki,Y.
TITLE     High-efficiency full-length cDNA cloning
JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
MEDLINE   99279253
PUBMED    10349636
AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE   20499374
PUBMED    11042159
AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,I., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE     RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE   20530913
PUBMED    11076861
AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE     Functional annotation of a full-length mouse cDNA collection
JOURNAL   Nature 409, 685-690 (2001)
REFERENCE 5

```

```

AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE     Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 2637)
AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE     Direct Submission
JOURNAL
COMMENT    Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]
COMMENT    cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
COMMENT    Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
COMMENT    Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
FEATURES
source    1..2637
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="NOD"
           /db_xref="FANTOM_DB:E430023D24"
           /db_xref="MGI:2427569"
           /db_xref="taxon:10090"
           /clone="E430023D24"
           /cell_type="thymic cells"
           /tissue_type="thymus"
           /clone_lib="RIKEN full-length enriched mouse cDNA library"
           /dev_stage="2 days neonate"
           358..1157
           /note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTR)AAL38196, evidence: FASTY, 100%ID, 100%length, match=777) putative"
           /codon_start=1
           /protein_id="BAC40495.1"
           /db_xref="GI:26353730"
           /translation="MGSLSRGKTSRSPSSGGQDPVSNQPRHKVTAVALGSFPA GEQRLURLGEPLTITISEDGDWTVQSEVSGREYHMPVSVYAKVAHWGLYEGLSREK AEELLPLNPGCGFLIRESCYRRCYSLSVLRSPASWDRIHYRIQRDNGLYLIS PRLTPPSLHALVEHYSELADGICCPLEPCVLQKLGPLPGKDTPPPTVPTSSLNWK LDRSLFLPEAPASGEASLSGLRESLSYSLSLAEDPLDDA"
CDS
ORIGIN
Alignment Scores:
Pred. No.:      8,96e-93      Length:      2637
Score:          1028.00      Matches:    208
Percent Similarity: 85.50%      Conservative: 16
Best Local Similarity: 79.39%      Mismatches: 34
Query Match:    75.98%      Indels:     4
DB:             11          Gaps:      3

```



```

/db_xref="taxon:10090"
/clone="5830437K10"
/sex="male"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
232..1011
/notes="unmaned protein product; MODULATOR OF ANTIGEN
RECEPTOR SIGNALING MARS (SPTR)AAL38196, evidence: FASTY,
100%ID, 100%length, match=777"
putative"
/codon_start=1
/protein_id="BAC27168.1"
/db_xref="GI:26326849"
/translation="MGLSRGKTSFSPSSSGPDQEPVSMQPERKHTAVALGSFPA
GEQARLSRGEDUTLISEGDWTVQSEVSGREYVYAKVAHGLWYGLGKREK
AEELLNPGFAGLIRSEQDTRGYSLVSRPDRINRHYIQRLDNGWLYIT
PRITFSLHALVHYSLAGICCPLEPCVLKGLPKDTPFPVTPVTSLNWKK
LDRSLLEAPASGEASLLSEGLRESLSYISLAEDPLDDA"
ORIGIN
Alignment Scores:
Pred. No.: 2,8e-92 Length: 2974
Score: 1024.00 Matches: 207
Percent Similarity: 85.50% Conservative: 17
Best Local Similarity: 79.01% Mismatches: 34
Query Match: 75.68% Indels: 4
DB: 11 Gaps: 3
US-09-939-853a-75 (1-261) x AK030877 (1-2974)
QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
DB 232 ATGGGAAGTTGTCCAGCAGAGGGAACCC---TCCAGCCCCAGCCGCTCTCTGGT 288
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGlnArgSerLysAlaThrAlaValLeu 40
DB 289 CCAGACAGGAACCCGTGTCTATCCACACCAAGAGACACAGGTCTCAGCTGGGCCCTG 348
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB 349 GCGAGTTTCCAGCAGGTGAACAGCCAGACTATCTCTGAGACTCGGGGAGCGCTGACC 408
QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGluTyr 80
DB 409 ATCATCTCTGAGGTAGAGATTGTGTGACAGTCCAGTCGGAAGTCTCAGCAGAGAGTAC 468
QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
DB 469 CACATGCCAGTGTGTATGTGGCTAAAGTCGCCACCGGTGGCTGTACGAGGCGCTGAGC 528
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120
DB 529 CGGAGAAAGCCGAGGAACTACTCTCTGTACCTGGGAACCCGAGGGGCTTCTCTCATC 588
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 589 CGGAGAGCCAGACAGAGAGGTACTATTCTCTGTCCGTCCGACTCAGCCGCCCTGCA 648
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle 160
DB 649 TCTTGGGACCGGATCAGACACTCAGGATACAGCGTCTTGACAATGCTGGCTGTATCATC 708
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
DB 709 ACACCTGCGCTTACCTTCCCTCTACTCAGCCCTGTGTGGAGCATTTACTCTGAGTAGCA 768
QY 181 AspAspIleCysLeuLeuGlyGluProCysValLeuGlnArgAlaGlyProLeuPro 200
DB 769 GATGCGATCTGCTGCTCCCTCAGGAGGCGGTGTCTCTCGAGAGCTTGGGCCACTACCT 828
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220

```

```

Db 829 GCGAAGATACACCTCCACCTGTGCTGTGCGCAACATCATCATCTAAATGGAAGCTG 888
QY 221 AspSerSerLeuLeuPheSerGluAla---AlaThrGlyGluGluSerLeuSerGlu 239
Db 889 GACCGCACCTCTCTCTGTTTCTGGAAGCAGCTCGAGTGGGAGGCACTCTCTCTCAGTCA 948
QY 240 GlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 949 GCGCTCCGAGAGTCCCTCAGTCTCTATCATCAGCTGCTGGCTGAGGAC-----CCCTTGAT 1002
QY 260 AspAla 261
Db 1003 GATGCT 1008
RESULT 6
AK020837
LOCUS
DEFINITION
Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone:A930009E21 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING
MARS, full insert sequence.
ACCESSION
AK020837
VERSION
AK020837.1 GI:12861542
SOURCE
HTC; CAP trapper.
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Masumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
REFERENCE
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 926)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiracka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,

```





QY 260 Asp 260  
 Db 646 GAC 648

RESULT 8  
 LOCUS AL541041 1201 bp mRNA linear EST 12-MAY-2003  
 DEFINITION AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YK23  
 5-PRIME, mRNA sequence.

ACCESSION AL541041  
 VERSION AL541041.2 GI:30544829  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1201)  
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 15, 2001 this sequence version replaced gi:12871733.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 9825.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DE005AF12QF1&cluster=9825.r. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DE005AF12QF1.  
 Location/Qualifiers  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE005YK23"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with NotI and cloned into  
 the NotI and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

FEATURES  
 source

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.32e-77 Length: 1201  
 Score: 867.00 Matches: 168  
 Percent Similarity: 95.45% Conservative: 0  
 Best Local Similarity: 95.45% Mismatches: 8  
 Query Match: 64.08% Indels: 0  
 DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x AL541041 (1-1201)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20  
 Db 459 ATGGAGTCTGCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTGAGTCTCTGTC 518  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 519 CAAGCCAGGACCTGTGACCATCGAAGCAGAGAGAGCCAGCGGTGGCCCTG 578  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 579 GCGAGTTCCTCCGCGAGGTGGCCCGCCGAGTGTCTGACCTCGGAGCCATTGACC 638  
 QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyValGlyTyr 80

Db 639 ATCGTCTCTGAGGATGGAGACTGGTGGACGGTGTGTCTGAAGTCTCAGGCAGAGATAT 698  
 QY 81 AsnIleProSerValHisValGlyIysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100  
 Db 699 AACATCCCGCAGCGTCCAGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGCGCTGAGC 758  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeuIle 120  
 Db 759 AGGGAGAAAGCAGAGAACTGCTGTTGTACTCGGAACCCCTGGAGGGGCCCTTCCTCATC 818  
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140  
 Db 819 CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTCTGTGATCCGCTCAGCGGCCCTGCA 878  
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160  
 Db 879 TCTTGGACCGGATCAGACATCAGATCAGATCAGATCAGATCAGATCAGATCAGATC 938  
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyr 176  
 Db 939 TCACGGCGCTTACCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 986

RESULT 9  
 LOCUS BG284179 566 bp mRNA linear EST 21-FEB-2001  
 DEFINITION 602408226P1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4520382 5',  
 mRNA sequence.

ACCESSION BG284179  
 VERSION BG284179.1 GI:13034866  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 566)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs@mail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10418 row: c column: 07  
 High quality sequence start: 2  
 High quality sequence stop: 566.  
 Location/Qualifiers  
 1..566  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4520382"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="NIH MGC 91"  
 /clone\_lib="NIH MGC 91"  
 /note="Organ: Prostate; Vector: pCMV-SPORT6; Site: 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dr primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.04e-61 Length: 566  
 Score: 711.00 Matches: 136  
 Percent Similarity: 99.27% Conservative: 0  
 Best Local Similarity: 99.27% Mismatches: 1  
 Query Match: 52.55% Indels: 0  
 DB: 12 Gaps: 0

US-09-939-853A-75 (1-261) x BG284179 (1-566)

QY 125 ThrArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144  
 Db 4 ACAGCGTCCGGGCTCTTACTCTGTGAGTCCGCTCAGCGCCCTGCATCTGGACCGG 63  
 QY 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeu 164  
 Db 64 ATCAGACATACAGATCCACTGGCTTGACATGGCTGGCTGTACATCTCACCAGCCCTC 123  
 QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAlaCys 184  
 Db 124 ACCTTCCCTCAGTCCAGGCGCTGTGGACCACTTACTCTGAGCTGGCGATGACATCTGC 183  
 QY 185 CysLeuLeuGlyGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
 Db 184 TGCCTACTCAGAGGCGCTGTCTCTCAGAGGCTGGCCGCTCCCTGGCAGGATATA 243  
 QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224  
 Db 244 CCCCTACTGTGACTGTGCAGAGGACCACTCACTGGAAGAGCTGGACACTCCCTC 303  
 QY 225 LeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGluGlyLeuArgGluSer 244  
 Db 304 CTGTTTCTGAAGCTGCCACAGGGAGGAGTCTTCTCAGTGGGCTCTCCGGGAGTCC 363  
 QY 245 LeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAla 261  
 Db 364 CTCAGCTTCTACATCAGCCTGATGACGAGGCTGTCTTTGGATGATGCC 414

RESULT 10  
 BQ054265  
 LOCUS  
 DEFINITION AGNCOURT 6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339  
 5', mRNA sequence.  
 ACCESSION BQ054265  
 VERSION BQ054265.1 GI:19813605  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 JOURNAL 1 (bases 1 to 986)  
 COMMENT NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM2125 row: i column: 12  
 High quality sequence stop: 515.

FEATURES  
 source  
 1..986  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5936339"  
 /tissue\_types="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_106"  
 /notes="Organ: blood; Vector: pOT7; Site: 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN

Alignment Scores:  
 Pred. No.: 3,35e-58 Length: 986  
 Score: 680.50 Matches: 163  
 Percent Similarity: 73.33% Conservative: 3  
 Best Local Similarity: 71.81% Mismatches: 17  
 Query Match: 50.30% Indels: 44  
 DB: 12 Gaps: 3

US-09-939-853A-75 (1-261) x BQ054265 (1-986)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20  
 Db 279 ATGGGAAGTCTGCCAGCAGAGAATAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 338  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 Db 339 CAAGCCAGGACCTGTGACCATGCAAGCAGAGAGAGCCACAGCCGTGGCCCTG 398  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 399 GGCAGTCTCCCGCAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 458  
 QY 61 IleValSerGluAspGlyVasPTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 459 ATGCTCTCTGAGGATGGAGACTGTGGACGGTGTCTGTGAGTCTCAGGAGAGATAT 518  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 Db 519 AACATCCCGCAGCTCCACGTCGCCCAAAAGTCTCCATGGTGGCTGTATGAGGGCTGAGC 578  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAla-PheLeuIle 120  
 Db 579 AGGAGAAAGCAGAGGAACTGCTGTTGTACTGGGACCCCTGAGGGGCCCTTCTCAT 638  
 QY 120 eArgGlu-SerGlnThrArgArgGlySerTyrSerLeuSerVal-ArgLeuSerArg-Pr 139  
 Db 639 CCGGAGAAAGCAGACAGGAGAGGCTTACTCTCTGCTCAGTCCCGCTCAGCGGCC 698  
 QY 139 OAlaSerTrp-AspArg-IleArgHisTyrArg-IleHisCysLeuAspAsnGlyTrpLe 158  
 Db 699 TGCACTCTGGGACCGGATCAGACCCCTACAGGATTCCTCCCTGCGCTTGAACCATTTGCT 758  
 QY 158 u-----TyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValas 174  
 Db 759 TGGCTTGTAAATTTTAAACCGGGGCTTACCCTTTCCC----- 798  
 QY 174 PHisTyrSerGluLeuAlaAspAspIleCysLeuLeuLysGluProCysValLeuGlu 194  
 Db 799 -----CTTAA 803  
 QY 194 nArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPr 214  
 Db 804 ATTTTCAGGGGCCCTTGGGGGAAACCCCATTT-----ACTCC 839  
 QY 214 OleuAsn 216  
 Db 840 TTAAAC 846

RESULT 11

LOCUS BQ944126  
 DEFINITION AGNCOURT 10545003 NIH\_MGC\_107 Homo sapiens cDNA clone  
 IMAGE:6728350 5', mRNA sequence.  
 ACCESSION BQ944126  
 VERSION BQ944126  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 BQ944126  
 AGNCOURT 10545003 NIH\_MGC\_107 Homo sapiens cDNA clone  
 IMAGE:6728350 5', mRNA sequence.  
 GI:24132945  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLCM3049 row: m column: 21
              High quality sequence stop: 628.
FEATURES
source
1..960
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:6728350"
   /tissue_type="adenocarcinoma, cell line"
   /lab_host="DH10B (phage-resistant)"
   /clone_lib="NIH_MGC_107"
   /note="Organ: Breast; Vector: pOTB7; Site: 1: EcoRI;
   Site: 2: XhoI; cDNA made by oligo-dt priming.
   Directionally cloned into EcoRI/XhoI sites using the
   following 5' adaptor: GGCACGAG(G). Library constructed by
   Ling Hong in the laboratory of Gerald M. Rubin (University
   of California, Berkeley) using ZAP-cDNA synthesis kit
   (Stratagene) and Superscript II RT (Life Technologies).
   Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      3.27e-51      Length:      960
Score:          611.50       Matches:     124
Percent Similarity: 86.71%    Conservative: 0
Best Local Similarity: 86.71% Mismatches:     1
Query Match:    45.20%       Indels:      18
DB:             13          Gaps:        1

US-09-939-853A-75 (1-261) x BU944126 (1-960)

Qy 120 lIeArgGlueRglnThrArgArgGlySerTySerLeuSerValArgLeuSerArgPro 139
Db 3 ATCCGGAGAGCCACAGCAGGAGGCTTACTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 62

Qy 140 AlaSerTrpAspArgIleArgHisTyArgIleHisCysLeuAspAsnGlyTrpLeuTyr 159
Db 63 GCATCTCTGGGACCGGATCAGACACTACAGATCCACTGCTTGCATATGGCTGGCTGTAC 122

Qy 160 lIeSerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTySerGluLeu 179
Db 123 ATCTCAGCGGCGCTCAGCTTCCCTCACTCCAGGCGCTGTGGACCATTC----- 173

Qy 180 AlaAspAspIleCysCysLeuLeuGluProCysValLeuGln-ArgAlaGlyProLe 199
Db 174 -----TCTGAGGCGCTGGCCCGCT 191

Qy 199 uProGlyLeuAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpIysG 219
Db 192 CCCTGGCAAGGATATACCCCTACCTGTGACTGTGCAGAGGACACCACTCACTGGAAGA 251

Qy 219 uLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerG 239
Db 252 GCTGGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGGAGGCTCTCTCTAGTGA 311

Qy 239 uGlyLeuArgGluSerLeuSerPheTyrlIeSerLeuAsnAspGluAlaValSerLeu 259
Db 312 GGGTCTCCGGAGTCCCTCAGCTTCTACATCAGCTGATGACGAGGCTGTCTCTTTGGA 371

Qy 259 pAspAla 261
Db 372 TGATGCC 378

RESULT 12
LOCUS    BQ053486
DEFINITION BQ053486
          5', mRNA sequence.
ACCESSION BQ053486
VERSION   BQ053486.1 GI:19812826
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 878)
           NIH-MGC http://mgs.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
           cDNA Library Preparation: Rubin Laboratory
           cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM2122 row: l column: 06
           High quality sequence stop: 394.
           Location/Qualifiers
             1..878
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5935253"
               /tissue_type="natural killer cells, cell line"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH_MGC_106"
               /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
               EcoRI; cDNA made by oligo-dt priming. Directionally cloned
               into EcoRI/XhoI sites using the following 5' adaptor:
               GGCACGAG(G). Library constructed by Ling Hong in the
               laboratory of Gerald M. Rubin (University of California,
               Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
               Superscript II RT (Life Technologies). Note: this is a
               NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      4.25e-46      Length:      878
Score:          560.50       Matches:     134
Percent Similarity: 63.60%    Conservative: 11
Best Local Similarity: 58.77% Mismatches:     21
Query Match:    41.43%       Indels:      64
DB:             12          Gaps:        5

US-09-939-853A-75 (1-261) x BQ053486 (1-878)

Qy 1 MetGlySerLeuProSerArgArgIysSerLeuProSerProSerLeuSerSerVal 20
Db 273 ATGGGAAGTGTGCCAGCAGAGAATAATCTCTGCCAGCCCAAGCTTGAGTTCCTCTGTC 332

Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40
Db 333 CAAGGCCAGGACCTGTGACCATGGAAGACAGAGAAGCAAGGCCACGCGTGGCCCTG 392

Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 393 CGACGTTTCCCGCAGGAGTGGCCCGCGAGGTGTCGTGAGACTCGGGAGCCATTGACC 452

Qy 61 lleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80

```



Best Local Similarity: 80.85% Mismatches: 18  
Query Match: 41.13% Incls: 1  
DB: 10 Gaps: 1

US-09-939-853A-75 (1-261) x BB635615 (1-660)

QY	1	MetGlySerLeuProSerArgArgLySSerLeuProSerProSerLeuSerSerVal	20
DB	241	ATGGGAAGTTTTCACGACAGAGGGAACCC---TCCAGGCCCCAGCCCCAGCTCCTCTGGT	297
QY	21	GingGlyGingLyProValThrMetGluAlaGluArgSerLyAlaThrAlaValAlaLeu	40
DB	298	CCAGACCAGGAACCCGTGTCCATGCAACACAGAAAGACACAGGTTCACAGCTGTGGCCCTG	357
QY	41	GlySerPreProAlaGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr	60
DB	358	GGCAGGTTTCCCAAGCAGGTGAACAGGCCAGACATCTCTGAGACTCGGGAGCCGCTGACC	417
QY	61	IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr	80
DB	418	ATCATCTCTGAGGATGGAGATGGTGACAGTCCAGTCGGAAGTCTCAGGCAGAGATAC	477
QY	81	AsnIleProSerValHisValGlyLysValSerHisGlyTyrIleuTyrGluGlyLeuSer	100
DB	478	CACATGCCACAGTGTGTATGTGGCTAAAGTCGCCACCGGTGGCTGTACGAGGCCCTGAGC	537
QY	101	ArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyAlaPheLeuIle	120
DB	538	CGGAGAAACCGCAGGAATCTCTCTGTTACCTGTGAAACCCCGAGGGGCTTCTTCATC	597
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
DB	598	CGGAGAGCCAGACACGAGGAGGCTGTATTCCCTGCTCGTTGACATCAGCCGCCCTGCA	657
QY	141	Ser 141	
DB	658	TCI 660	

AL844311 pool\_YT\_lib\_v\_SPD Homo sapiens linear EST 30-JUL-2002

AL844311

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: humquery@sanger.ac.uk  
Sanger Centre name: sccdi0818.154136A  
Homo sapiens EST sequence. This sequence was generated as part of  
The Wellcome Trust Sanger Institute program to identify and  
annotate genes in the human genome. Incomplete or unconfirmed genes  
are experimentally analysed using a variety of cDNA library  
resources. This sequence was obtained from a PCR product generated  
from a pool of up to 100,000 cDNA clones derived from  
pool\_YT lib v\_SPD cDNA library. Further information can be found at  
http://www.sanger.ac.uk/Teams/Team69/  
Location/Qualifiers  
1..597  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="20"

FEATURES

Source

/clone\_lib="pool YT lib v SPD"

ORIGIN

Alignment Scores: Pred. No.: 7.98e-46 Length: 597  
Score: 555.00 Matches: 107  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 41.02% Indels: 0  
DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x AL844311 (1-597)

QY	155	AsnGlyT	TrpLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAsp	174
Db	13	AATGGCTGGGTGACATCTCACCACGGCTCACCTTCCCTCACTCCAGGCCCTGGTGGAC		72
QY	175	HistyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysValLeuGln		194
Db	73	CATTACTCTGAGCTGGGGATGACATCTGCTGCTACTCAAGGAGCCCTGTGTCTCTGCAG		132
QY	195	ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPro		214
Db	133	AGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACTGTGAGAGGACACCA		192
QY	215	LeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGlu		234
Db	193	CTCAACTGGAAAGAGCTGGACAGCTCCCTCTCTTTCTGAGCTGCCACAGGGAGGAG		252
QY	235	SerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu		254
Db	253	TCTCTTCTCAGTGAGGGTCTCCGGGAGTCCCTCAGCTTCTACATCATGCTGATGACGAG		312
QY	255	AlaValSerLeuAspAspAla		261
Db	313	GCTGCTCTTTGGATGATGCC		333

Search completed: February 20, 2004, 11:30:54  
Job time : 2531 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 09:40:13 ; Search time 39 Seconds  
(without alignments)  
1627.440 Million cell updates/sec

Title: US-09-939-853A-75  
Perfect score: 1353  
Sequence: 1 MSLPERRKSLPSPSSSV.....RESLSPYISLNDVAVSLDDA 261

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US09939853/tunat\_19022004\_24508/app\_query.fasta\_1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -DOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09939853 @CIGN 1.1.85 @tunat\_19022004\_24508 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	370.5	27.4	2298	4	US-09-023-655-1158
2	360.5	26.6	2015	4	Sequence 1158, Ap
3	340	25.1	2129	4	Sequence 1105, Ap
4	323	23.9	2435	4	Sequence 1452, Ap
5	320	23.7	2647	4	Sequence 1313, Ap
6	320	23.7	2647	5	Sequence 77, Appl
7	315.5	23.3	4517	4	PCT-US93-06251-77
8	315.5	23.3	4517	5	Sequence 7, Appl
9	313.5	23.2	1491	2	US-09-470-881-7
10	313.5	23.2	1491	2	PCT-US93-06251-83
11	312.5	23.1	2354	4	Sequence 83, Appl
12	289	21.4	1759	4	Sequence 1, Appl
					Sequence 1080, Ap
					Sequence 2, Appl

13	287	21.2	1602	1	US-07-820-011A-1
14	287	21.2	1602	5	PCT-US93-00445-1
15	276.5	20.4	1611	1	US-07-820-011A-3
16	276.5	20.4	1611	4	US-09-860-473-3
17	276.5	20.4	1611	5	PCT-US93-00445-3
18	273	20.2	1626	4	US-09-860-473-10
19	262	19.4	675	1	US-08-707-793A-3
20	262	19.4	675	1	US-08-707-792A-3
21	243.5	18.0	2827	4	US-08-492-723-1
22	240.5	17.8	2770	4	US-08-426-509A-5
23	240.5	17.8	2770	4	US-08-232-545-5
24	240.5	17.8	2770	5	PCT-US95-05008-5
25	240.5	17.8	2863	4	US-09-023-655-1389
26	240.5	17.8	7607	1	US-08-222-616-19
27	240.5	17.8	7607	4	US-08-446-648-19
28	240.5	17.8	7607	4	US-09-982-610-19
29	240.5	17.8	7607	5	PCT-US95-04228-19
30	238	17.6	282	2	US-09-006-675-5
31	238	17.6	282	3	US-09-238-603A-5
32	230	17.0	1467	4	US-09-579-182-2
33	230	17.0	1548	4	US-09-099-053-1
34	202	14.9	1661	2	US-08-815-176-2
35	202	14.9	1661	4	US-09-197-344-2
36	193.5	14.3	2187	4	US-09-023-655-1267
37	193.5	14.3	2187	4	US-09-470-881-4
38	190.5	14.1	3623	1	US-08-306-691B-35
39	182	13.5	1804	1	US-08-306-691B-40
40	182	13.5	1804	4	US-09-167-322-14
41	182	13.5	1804	5	PCT-US93-06251-82
42	172	12.7	874	4	US-09-023-655-931
43	172	12.7	1072	1	US-07-906-349A-2
44	172	12.7	1072	1	US-08-167-035-5
45	172	12.7	1072	1	US-08-167-035-49

#### ALIGNMENTS

#### RESULT 1

US-09-023-655-1158  
; Sequence 1158, Application US/09023655  
; Patent No. 6607979

; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: FA-0001 US  
; TELECOMMUNICATION INFORMATION:

```
;
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1158:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2298 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g187268
US-09-023-655-1158

Alignment Scores:
Pred. No.: 9.15e-30 Length: 2298
Score: 370.50 Matches: 80
Percent Similarity: 57.71% Conservative: 36
Best Local Similarity: 39.80% Mismatches: 76
Query Match: 27.38% Indels: 9
DB: 4 Gaps: 3

US-09-939-853A-75 (1-261) x US-09-023-655-1158 (1-2298)
QY 6 SerArgGlySerLeuProSerProSerLeuSerSerSerValGlnGlyPro 25
Db 409 TCCATAAACAGCAAGCCAGTTCAGAA---TCTCAGCTTTTACCTGGACAGAGTTT 465
QY 26 ValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerProAla 45
Db 466 CAACCTAAGATCCAGAGGACAGGACATCTGTGTAGCTTTGTACCCCTATGATGGC 525
QY 46 GlyGlyProAlaGluLeuSerLeuGlyGlyProLeuThrLeuValSerGluAsp 65
Db 526 ATCCACCGGACGACTTGTCTTCAAGAAAGGAGAGATGAAGTCTGTGGAGGACAT 585
QY 66 GlyAspTrpTptThrValLeuSerGluValSerGlyArgGluTyrAsnIleProSerVal 85
Db 586 CGAATGCTGGAAGCAAGTCTCTTTACAAAAAAGAGCTTCATCCCGACCAAC 645
QY 86 HisValGlyLysVal-----SerHisGlyTrpLeuTyrGluGlyLeuSerArg 101
Db 646 TATGTGGCCAACTCAACACCTTAGAAACAGAGAGTGGTTTTTCAAGGATATAACACGG 705
QY 102 GluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuIleArg 121
Db 706 AAGACGCGAAGAGCAGCTTTTGGCCAGGAATAGCGCTGGAGCTTTCTTTATTAGA 765
QY 122 GluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSer 141
Db 766 GAAAGTGAACATTAAAGGAAGCTTCTCTGTCTGTCTGAGAGACTTTGACCCCTGTGCAT 825
QY 142 TrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIleSer 161
Db 826 GGTGATGTTATTAGCACTACAAATATAGAAAGTCTGGATATGGGGCTATTACATCTCT 885
QY 162 ProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAsp 181
Db 886 CCACGATCACTTTTCCCTGTATCAGCGACATGATTAAACATTACCAAAAGCAGGAGAT 945
QY 182 AsprIleCysLysLeuLeuGlyGluProCysValLeuGlnArgAlaGlyProLeuProGly 201
Db 946 GGCCTTGTGCAGAAATTGGAGAAGCTTGTTATT-----AGTCCCAAGCCACAG 993
QY 202 Lys 202
Db 994 Arg 996

RESULT 2
US-09-023-655-1105
; Sequence 1105, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
```

APPLICANT: Susan G. Stuart  
APPLICANT: Jeffrey J. Sellhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023.655  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1105:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g183911  
US-09-023-655-1105

Alignment Scores:  
Pred. No.: 8.84e-29 Length: 2015  
Score: 360.50 Matches: 77  
Percent Similarity: 58.38% Conservative: 31  
Best Local Similarity: 41.62% Mismatches: 70  
Query Match: 26.64% Indels: 7  
DB: 4 Gaps: 2

US-09-939-853A-75 (1-261) x US-09-023-655-1105 (1-2015)

QY 12 ProSerProSerLeuSerSerSerValGlnGlyProValThrMetGluAlaGlu 31  
Db 286 CCGGGGCTTAATAGACACACAGC-----AACACACAGGAATCAGGAGGCGAGGC 336  
QY 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeu 51  
Db 337 TCTGAGGACATCATCTGTGTGCTCTGTATGATGAGGAGCCATTCACCCAGACGCTC 396  
QY 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrVal 71  
Db 397 AGCTTCCAGAGGGGACACAGATGTTGTCTTAGAGGAATCCGGGAGTGTGTGAAGGCT 456  
QY 72 LeuSerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal--- 90  
Db 457 CGATCTCTGGCCACCGGAGAGGAGGCTACATCCACAGCAACTATGTCCCGCGTTCAC 516  
QY 91 -----SerHisGlyTrpLeuTyrGluGlyLeuSerArgGluLysAlaGluLeu 107  
Db 517 TCTCTGGACAGAGAGTGTGTTTTTCAAGGCGCATCAGCCGAGAGGCGCAGAGCGCAA 576



NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1313:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2435 base pairs  
TYPE: nucleic acid  
STRADEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: 9338227  
US-09-023-655-1313

Alignment Scores:  
Pred. No.: 1,248-24 Length: 2435  
Score: 323.00 Matches: 81  
Percent Similarity: 51.74% Conservative: 23  
Best Local Similarity: 40.30% Mismatches: 76  
Query Match: 23.87% Indels: 22  
DB: 4 Gaps: 3

US-09-939-853A-75 (1-261) x US-09-023-655-1313 (1-2435)

```
QY 12 ProSerProSerLeuSerSerValGlnGly----- 22
DB 507 CCATCCCACTTACAACTTCCAGCGAGCGGGGCCAAGGACTCACCGTCTTTGGAG 566
QY 23 -----GlnGlyProValThrMetGluAlaGluArgSerLysAla 35
DB 567 GTGTGAACCTCTCGTCTCATACGGGACCTTCGCTACGAGAGGAGGAACAGAGTG-ACA 625
QY 36 ThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeu 55
DB 626 CTCCTTTGTGGCCCTTTATGACTATGACGAGCGGACGAGAGATGACCTGAGTTTCAAAA 685
QY 56 GlyGluProLeuThrIleVal---SerGluAspGlyAspTTPTrpThrValLeuSerGlu 74
DB 686 GGAGAAAAATTTCAAATATTGAACAGCTCGGAGAGGAGATTGTTGGGAGCCCGCTCTTG 745
QY 75 ValSerGlyArgGluTyAsnIleProSerValHisValGlyLysVal----- 90
DB 746 ACAACTGGAGAGACAGGTTTACATTTCCAGCAATTATGTGGCTCCAGTTGACTCTATCCAG 805
QY 91 SerHisGlyTTPLeuTyGluLeuSerArgGluLysAlaGluGluLeuLeuLeu 110
DB 806 GCAGAGAGTGTACTTTGGAAAACCTTGGCCGAAAAGATGCTGAGCGACACTATTGCTCC 865
QY 111 ProGlyAsnProGlyGlyAlaPheLeuLeuLeuLeuLeuLeuLeuLeuLeu 130
DB 866 TTTGGAAAACCCCAAGAGGTACCTTTCTTATCCGCGAGAGTGAACCAACCAAGGTCCTAT 925
QY 131 SerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgile 150
DB 926 TCACITTTCTACCGTCATGGATGATATGAAGAGACCATGTCAAACATTATAAAATT 985
QY 151 HisCysLeuAspAsnGlyTTPLeuTyTTPLeuTyTTPLeuTyTTPLeuTyTTPLeuTy 170
DB 986 CGCAAACTTGACATGTTGGATGATATGATATGATATGATATGATATGATATGATATGATAT 1045
QY 171 AlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluPro 190
DB 1046 CAGCTTTGTACAACATTACTTCAGAGAGAGCTGCGAGTCTCTGCTGCGCCTAGTAGTCCC 1105
QY 191 Cys 191
DB 1106 TGT 1108
```

RESULT 5

US-09-220-132-77

Sequence 77, Application US/09220132  
Patent No. 6506607  
GENERAL INFORMATION:

APPLICANT: Shvjan, Andrew W.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 07334-074001  
CURRENT APPLICATION NUMBER: US/09/220,132  
CURRENT FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: US 60/079,303  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: US 60/069,821  
PRIOR FILING DATE: 1997-12-24  
NUMBER OF SEQ ID NOS: 191  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 77  
LENGTH: 2647  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-220-132-77

Alignment Scores:  
Pred. No.: 2,978-24 Length: 2647  
Score: 320.00 Matches: 80  
Percent Similarity: 51.74% Conservative: 24  
Best Local Similarity: 39.80% Mismatches: 76  
Query Match: 23.65% Indels: 22  
DB: 4 Gaps: 3

US-09-939-853A-75 (1-261) x US-09-220-132-77 (1-2647)

```
QY 12 ProSerProSerLeuSerSerValGlnGly----- 22
DB 716 CCATCCCACTTACAACTTCCAGCGAGCGGGGCCAAGGACTCACCGTCTTTGGAG 775
QY 23 -----GlnGlyProValThrMetGluAlaGluArgSerLysAla 35
DB 776 GTGTGAACCTCTCGTCTCATACGGGACCTTCGCTACGAGAGGAGGAACAGAGTG-ACA 834
QY 36 ThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeu 55
DB 835 CTCCTTTGTGGCCCTTTATGACTATGACGAGCGGACGAGAGATGACCTGAGTTTCAAAA 894
QY 56 GlyGluProLeuThrIleVal---SerGluAspGlyAspTTPTrpThrValLeuSerGlu 74
DB 895 GGAGAAAAATTTCAAATATTGAACAGCTCGGAGAGGAGATTGTTGGGAGCCCGCTCTTG 954
QY 75 ValSerGlyArgGluTyAsnIleProSerValHisValGlyLysVal----- 90
DB 955 ACAACTGGAGAGACAGGTTTACATTTCCAGCAATTATGTGGCTCCAGTTGACTCTATCCAG 1014
QY 91 SerHisGlyTTPLeuTyGluLeuSerArgGluLysAlaGluLeuLeuLeuLeu 110
DB 1015 GCAGAGAGTGTACTTTGGAAAACCTTGGCCGAAAAGATGCTGAGCGACAGCTATTGTC 1074
QY 111 ProGlyAsnProGlyGlyAlaPheLeuLeuLeuLeuLeuLeuLeuLeuLeu 130
DB 1075 TTTGGAAAACCCCAAGAGGTACCTTTCTTATCCGCGAGAGTGAACCAACCAAGTCCCTAT 1134
QY 131 SerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgile 150
DB 1135 TCACITTTCTACCGTCATGGATGATATGATATGATATGATATGATATGATATGATATGATAT 1194
QY 151 HisCysLeuAspAsnGlyTTPLeuTyTTPLeuTyTTPLeuTyTTPLeuTyTTPLeuTy 170
DB 1195 CGCAAACTTGACATGTTGGATGATATGATATGATATGATATGATATGATATGATATGATAT 1254
QY 171 AlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluPro 190
DB 1255 CAGCTTTGTACAACATTACTTCAGAGAGAGCTGCGAGTCTCTGCTGCGCCTAGTAGTCCC 1314
QY 191 Cys 191
DB 1191 Cys 191
```

Db 1315 TGT 1317

RESULT 6

PCT-US93-06251-77

Sequence 77, Application PC/TUS9306251

GENERAL INFORMATION:

APPLICANT: Wikstrom, Eric and Rife, Jason P.

TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing Stereospecific Alkylphosphonates and Arylphosphonates

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: NY

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/06251

FILING DATE: 19930630

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8586

TELEPHONE: 516-742-4343

TELEFAX: 516-742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 77:

SEQUENCE CHARACTERISTICS:

LENGTH: 2647 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US93-06251-77

Alignment Scores:

Pred. No.: 2,97e-24 Length: 2647

Score: 320.00 Matches: 80

Percent Similarity: 51.74% Conservative: 24

Best Local Similarity: 39.80% Mismatches: 76

Query Match: 23.65% Indels: 22

DB: 5 Gaps: 3

US-09-939-853A-75 (1-261) x PCT-US93-06251-77 (1-2647)

Qy 12 ProSerProSerLeuSerSerValGlnGly----- 22

Db 716 CCATCCCACTACAACTTCCAGCGCGGGCCAGGAGCTCACCCTTTGGAG 775

Qy 23 -----GlnGlyProValThrMetGluAlaGluArgSerLysala 35

Db 776 GTGTGAAGTCTTCGTCTCATCGGAGACCTTCGTACGAGAGGAGGAGACAGAGGTG-ACA 834

Qy 36 ThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeu 55

Db 835 CTCTTTGTGCCCCCTTTATGACTATGAGCAGCGAGAGAGATGACCTGATTTTACAAA 894

Qy 56 GlyGluProLeuThrIleVal-----SerGluAspGlyAspTyrThrValLeuSerGlu 74

Db 895 GGAGAAAAATTTCAAATATTGAACAGCTCGGAGGAGATTGCTGGAGACCGGCTCTTG 954

Qy 75 ValSerGlyArgGluTyrAsnIleProSerValHisValGlyVal----- 90

Db 955 ACAACTGGAGACAGAGTTACATTCAGCAATTATGTGGTCCAGCTTACTTATCCAG 1014

Qy 91 SerHisGlyTrpLeuTyrGluGlyLeuSerArgGluLysAlaGluGluLeuLeuLeu 110

Db 1015 GCAGAAGAGTGGTACTTTGGAAACTTGGCCGAAAGATGCTGAGCGACAGCTATTGTCC 1074

Qy 111 ProGlyAsnProGlyGlyAlaPheLeuLeuLeuSerGlnThrArgA-GlySerTyr 130

Db 1075 TTTGGAACCCAGAGGTACCTTTCTATCCGAGAGTGAACCCACCAAGGTGCTAT 1134

Qy 131 SerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgile 150

Db 1135 TCACTTTCTATCCGTGATTGGATGATGAAAGGAGACCATGTCAAACATTATAAAATT 1194

Qy 151 HisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeuThrPheProSerLeuGln 170

Db 1195 CGCAACTTGACAATGTGTGACTACTACATCACCCGCGGCCAGTTTGAACACACTTCAG 1254

Qy 171 AlaLeuValAspHisTyrSerGluLeuAlaAspIleCysCysLeuLysGluPro 190

Db 1255 CAGCTTGTACAACTTACTCAGAGAGAGCTGCAGGTCTCTGCTGCCGCTAGTAGTTCCC 1314

Qy 191 Cys 191

Db 1315 TGT 1317

RESULT 7

US-09-470-881-7

Sequence 7, Application US/09470881

Patent No. 6685938

GENERAL INFORMATION:

APPLICANT: CHERESH, David A.

APPLICANT: ELICEIRI, Brian

TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR MODULATION OF ANGIOGENESIS AND VASCULAR PERMEABILITY USING SRC OR YES TYROSINE KINASES

TITLE OF INVENTION: YES TYROSINE KINASES

FILE REFERENCE: TSRI 651.2

CURRENT APPLICATION NUMBER: US/09/470,881

PRIOR FILING DATE: 1999-12-22

PRIOR APPLICATION NUMBER: PCT/US99/11780

PRIOR FILING DATE: 1999-05-28

PRIOR APPLICATION NUMBER: 60/087,220

PRIOR FILING DATE: 1998-05-29

NUMBER OF SEQ ID NOS: 8

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 7

LENGTH: 4517

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (208)..(1836)

OTHER INFORMATION: human Yes-1 cDNA translated protein

US-09-470-881-7

Alignment Scores:

Pred. No.: 2,05e-23 Length: 4517

Score: 315.50 Matches: 92

Percent Similarity: 43.22% Conservative: 45

Best Local Similarity: 29.02% Mismatches: 113

Query Match: 23.32% Indels: 67

DB: 4 Gaps: 9

US-09-939-853A-75 (1-261) x US-09-470-881-7 (1-4517)

Qy 2 GlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal--- 20

Db 304 GGAGCAGACCCACTACAGTGTCCACCATGTCCGTCTATCTTCAGCAAGGAGGACAGCAGTT 363

Qy 21 -----GlnGlyGlnGlyProValThrMetGluAla 30

Db 364 AATTTCAGCAGTCTTTCCATGACACCATTTGGAGATCTCAGGGGTAAACGCTTTTGA 423

Qy 31 GluArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGly----- 46

```
Db 424 GGTGCATCTTCTCATTTTCAGTGTGCGCAAGTTTCATATCTCTGCTGTTTAAACAGTGGT 483
Qy 47 -----GlyProAlaGluLeuSerLeu 53
Db 484 GTTACTATATTGTGGCTTATATCATTTATGAGCTAGAACTACAGAGACCTTTCATTT 543
Qy 54 ArgLeuGlyGluProLeuThrIleValSerGlu---AspGlyAspTyrThrValLeu 72
Db 544 AAGAAGGTTGAAGATTTCAAATATTAACAATACGAAGAGAGATTGGTGGGAACAAGA 603
Qy 73 SerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal----- 90
Db 604 TCAATCGCTACAGGAAGAAGTGTATATCCGAGCAATTATAGCGCTGCAGATCC 663
Qy 91 -----SerHisGlyTyrLeuTyrGluGlyLeuSerArgGluLysAlaGluGluLeu 108
Db 664 ATTCAGCGAGAAGATGGTATTTCGCAAAATGGGGAAGAAAGATCTCGAAAGATTACT 723
Qy 109 LeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgGly 128
Db 724 TTGAATCTGGAAATCAACGAGGATTTCCTTAGTAGAGAGAGTGAACACTAAAGGT 783
Qy 129 SerTyrSerLeuSerValArgLeuSerArgProAlaSerTyrAspArgIleArg----- 146
Db 784 GCTTATTCCTTTCATTTCGT-----GATTGGATGAGTAAAGGGTGAC 828
Qy 147 -----HisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIleSerProArg 163
Db 829 AATGTGAACACTACAAATATAGAAACCTTGCAATGGTGGATACTATATCACCAACAGA 888
Qy 164 LeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspIle 183
Db 889 GCACAATTTGATCTCTGCAGAAATTTGGTGAACACTACACAGAACATGCTGATGGTTA 948
Qy 184 CysCysLeuLeuLysGluProCys-----ValLeuGln 194
Db 949 TCCACAGATTGCACAACTGTGTGCCAAGTGTGAACCTCAGACTCAAGGTCTAGCAAA 1008
Qy 195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArg----- 212
Db 1009 GATGCTGGGAATCCCTCGAAGATCTTGGCACTAGAGGTTAAACTAGGACAGGATGT 1068
Qy 213 -----ThrProLeuAsnTyrLysGluLeu 220
Db 1069 TTCGGCGAAGTGTGATGGGAACATGGAATGGAAACCCAGCAAGTAGCAATCAAAACACTA 1128
Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
Db 1129 AAACAGGTACAAATGATCCAGAACTTTCCTCAAGAAGCTCAGATAATGAAAAATTA 1188
Qy 241 LeuArgGluSerLeu---SerPheTyrIleSerLeuAsnAspGluAlaVal 256
Db 1189 AGACATGATAAACTTGTTCCTCATATATGCTGTGTTCCTGAAGAACCAATT 1239
```

## RESULT 8

PCT-US93-06251-83

Sequence 83, Application PC/TUS9306251

GENERAL INFORMATION:

APPLICANT: Wickstrom, Eric and Rife, Jason P.

TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing

TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESSES: SCULLY, SCOTT, MURPHY &amp; PRESSER

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: NY

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/06251

FILING DATE: 19930630

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8586

TELECOMMUNICATION INFORMATION:

TELEPHONE: 516-742-4343

TELEFAX: 516-742-4366

TELEX: 230 901 SAMS UR

INFORMATION FOR SEQ ID NO: 83:

SEQUENCE CHARACTERISTICS:

LENGTH: 4517 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US93-06251-83

## Alignment Scores:

Pred. No.:	2,05e-23	Length:	4517
Score:	315.50	Matches:	92
Percent Similarity:	43.22%	Conservative:	45
Best Local Similarity:	29.02%	Mismatches:	113
Query Match:	23.32%	Indels:	67
DB:	5	Gaps:	9

US-09-939-853A-75 (1-261) x PCT-US93-06251-83 (1-4517)

```
Qy 2 GlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal---- 20
Db 304 GGAGCAGAACCCACATACAGTGCACCATGTCGTCATCTTCAGCAAGGGAACAGCAGTT 363
Qy 21 -----GlnGlyGlnGlyProValThrMetGluAla 30
Db 364 AATTTCAGCAGTCTTTCATGACACCATTTGGAGGATCCTCAGGGGTAAACGCTTTTGA 423
Qy 31 GluArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGly----- 46
Db 424 GGTGCATCTCTCTCATTTTCAGTGGTCCCAAGTTTCATATCTCTGCTGTTTAAAGGTGT 483
Qy 47 -----GlyProAlaGluLeuSerLeu 53
Db 484 GTTACTATATTTGTGGCTTATATGATGATATGAAGCTAGAACTACAGAACCTTCATTT 543
Qy 54 ArgLeuGlyGluProLeuThrIleValSerGlu---AspGlyAspTyrThrValLeu 72
Db 544 AAGAAGGTTGAAGATTTCAAATATTAACAATACGAAGAGAGATTGGTGGGAACAAGA 603
Qy 73 SerGluValSerGlyArgGluTyrAsnIleProSerValHisValGlyLysVal----- 90
Db 604 TCAATCGCTACAGGAAGAAGTGTATATCCGAGCAATTATGATAGCGCTGCAGATCC 663
Qy 91 -----SerHisGlyTyrLeuTyrGluGlyLeuSerArgGluLysAlaGluGluLeu 108
Db 664 ATTCAGCGAGAAGATGGTATTTCGCAAAATGGGGAAGAAAGATCTCGAAAGATTACT 723
Qy 109 LeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgGly 128
Db 724 TTGAATCTCGAAATCAACGAGGATTTCCTTAGTAGAGAGAGTGAACACTAAAGGT 783
Qy 129 SerTyrSerLeuSerValArgLeuSerArgProAlaSerTyrAspArgIleArg----- 146
Db 784 GCTTATTCCTTTCATTTCGT-----GATTGGATGAGTAAAGGGTGAC 828
Qy 147 -----HisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIleSerProArg 163
Db 829 AATGTGAACACTACAAATATAGAAACCTTGCAATGGTGGATACTATATCACCAACAGA 888
Qy 164 LeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspIle 183
```

Db 889 GCACATTTGATCTCTGCAGAAATTTGGTGAACACTACACAGAACATCTGATGTTTVA 948  
 Qy 184 CysCysLeuLeuYsgluProCys-----ValLeuGln 194  
 Db 949 TGGCACAAGTTGACAACTGTGTCCCACTGTGAAACCTCAGACTCAAGGTCTAGCAAAA 1008  
 Qy 195 ArgAlaGlyProLeuProGlyLysAspPleProLeuProValThrValGlnArg----- 212  
 Db 1009 GATGCTTTGGGAAATCCCTCGAGAATTTTTCGAGTAGAGTTAAACTAGGACCAAGGATG 1068  
 Qy 213 -----ThrProLeuAsnTrpLysGluLeu 220  
 Db 1069 TTCGGCGAAGTGTGGATGGGAACATGGAATGGAACCAAGTAGCAATCAAAACACTA 1128  
 Qy 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240  
 Db 1129 AAACCGAGTACATGATGCGAGAAGTTTCCTTCAAGAAGCTCAGATATAATGAAAAATTA 1188  
 Qy 241 LeuArgGluSerLeu---SerPheTyrlleSerLeuAsnAspGluAlaVal 256  
 Db 1189 AGACATGATAAACTTGTTCACATATATGCTGTGTTTCTGAGAACCAATT 1239

RESULT 9

US-09-006-675-1  
 ; Sequence 1, Application US/09006675  
 ; Patent No. 5952213  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hemmati-Brivanlou, Ali  
 ; APPLICANT: Weinstein, Daniel C.  
 ; TITLE OF INVENTION: A NOVEL SRC-FAMILY KINASE AND METHODS OF  
 ; TITLE OF INVENTION: USE THEREOF  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Klauber & Jackson  
 ; STREET: 411 Hackensack Avenue, 4th Floor  
 ; CITY: Hackensack  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07601  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; FILING DATE: 13-JAN-1998  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Jackson Esq, David A.  
 ; REGISTRATION NUMBER: 26,742  
 ; REFERENCE/DOCKET NUMBER: 600-1-217  
 ; TELEPHONE: 201-487-5800  
 ; TELEFAX: 201-343-1684  
 ; TELEX: 133521  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1491 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHETICAL: NO  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..1491  
 ; US-09-006-675-1

Alignment Scores: 6.12e-24 Length: 1491  
 Pred. No.: 313.50 Matches: 78  
 Score:

Percent Similarity: 52.34% Conservativity: 34  
 Best Local Similarity: 36.45% Mismatches: 75  
 Query Match: 23.17% Indels: 27  
 DB: 2 Gaps: 6  
 US-09-939-853A-75 (1-261) x US-09-006-675-1 (1-1491)  
 Qy 1 MetGlySerLeuProSerArg-----ArgLysSerLeuProSerProSer 15  
 Db 1 ATGGGCTGCATCAAGTCAAGTCAAGGATCAAAATACGACTGGCAAAAGTCTGGGACCTCCCGAA 60  
 Qy 16 LeuSerSerValGlnGlyPro-----ValThrMet-----GluAlaGlu 31  
 Db 61 AGCACCAACCCCATTTATGTGAGGACCCACACTCTACAGTAACTATGACTAAACCTGAA 120  
 Qy 32 ArgSer-----LysAlaThrAlaValAlaLeuGly 41  
 Db 121 AGATCATCTTAAGCACCCAGAGAGAGGCAAGAGAGTGTCTGCTGGCTTTTGTAT 180  
 Qy 42 SerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIle 61  
 Db 181 GACTATGATGAGTCCACCTCGGGATCTGACTTTAGGAAAGGGGACCATCTCTCTCTA 240  
 Qy 62 ValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyAsn 81  
 Db 241 AAGAAAGAGTCAGGGGAGTGTGTGGAAGCATGTCTAATTTCCACTGTGTAAGAAGGCTTT 300  
 Qy 82 IleProSerValHisValGlyLysVal-----SerHisGlyTrpLeuTyGlu 97  
 Db 301 GTTCCAGTACTATGAGGTATTTCAATTCCTGGAATCTGAAGAGTGTACTTTTAA 360  
 Qy 98 GlyLeuSerArgGlyLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyAla 117  
 Db 361 GGCATGAGCGGGAAGGAGTGTGAAAGCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
 Qy 118 PheLeuIleArgGluSerGlnThrArgArgGlySerTySerLeuSerValArgLeuSer 137  
 Db 421 TTCAATGTCGAGAGTGTGAGACATGAAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCT 474  
 Qy 138 ArgProAlaSerTrpAspArgIleArgHisTyArgIleHisCysLeuAspAsnGlyTrp 157  
 Db 475 -----GACTCAGGGGACACTGTGAAACATTACAAAATTCGCACACTCGATGAGGT 528  
 Qy 158 LeuTyrlleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAlaAspHisTySer 177  
 Db 529 TTTTCATTTCTACACGATCCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 588  
 Qy 178 GluLeuAlaAspAspIleCysCysLeuLeuLysGluProCys 191  
 Db 589 GGTAAGTGGATGGCTGTGTCTCAGTGCCTTACATACCATGC 630

RESULT 10

US-09-228-603A-1  
 ; Sequence 1, Application US/09228603A  
 ; Patent No. 6291651  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hemmati-Brivanlou, Ali  
 ; APPLICANT: Weinstein, Daniel C.  
 ; TITLE OF INVENTION: A NOVEL SRC-FAMILY KINASE AND METHODS OF  
 ; TITLE OF INVENTION: USE THEREOF  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Klauber & Jackson  
 ; STREET: 411 Hackensack Avenue, 4th Floor  
 ; CITY: Hackensack  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07601  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/228,603A  
FILING DATE: 12-JAN-1999  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-217 N  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1491 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1491  
US-09-228-603A-1

Alignment Scores:  
Pred. No.: 6,12e-24 Length: 1491  
Score: 313.50 Matches: 78  
Percent Similarity: 52.34% Conservativity: 34  
Best Local Similarity: 36.45% Mismatches: 75  
Query Match: 23.17% Indels: 27  
DB: 3 Gaps: 6

US-09-939-853A-75 (1-261) x US-09-228-603A-1 (1-1491)

```
QY 1 MetGlySerLeuProSerArg-----ArgLysSerLeuProSerProSer 15
DB 1 ATGGGCTGCTCAAGTCAAGTCAAGATCAATACGACTGCGCAAAAGTCTGGACCTCCGGA 60
QY 16 LeuSerSerValGlnGlyGlnGlyPro-----ValThrMet---GluAlaGlu 31
DB 61 AGCACCAACCCATTTATGTGAAGGACCCCAACATCTACGTAAGTACTAGCTAACTGAACTGAA 120
QY 32 ArgSer-----LysAlaThrAlaValAlaLeuGly 41
DB 121 AGATCATCTAAGCCCCAGAGAGAGGAGGCAAGAGAGTGTCTGCTGGCTTTGTAT 180
QY 42 SerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIle 61
DB 181 GACTATGATGGAGTCCACCTCGGGATCTGACTTTTAGGAAAGGGAGACCATCTCTGCTA 240
QY 62 ValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGlyTyrAsn 81
DB 241 AAGAAAGAGTCAGGGGAGTGTGTGAAGAGCATGTCTAATTTCCACTGTGTGAAGAGGCTTT 300
QY 82 IleProSerValHisValGlyLysVal-----SerHisGlyTyrLeuTyrGlu 97
DB 301 GTTCCCGAGTAATGATGAGGTATTTCAATCCCTGGAATCTGAAGAGTGTACTTTAA 360
QY 98 GlyLeuSerArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyVala 117
DB 361 GGCATGAGCGGAGGAGGAGTGAAGGAGTGTATCTCTGTTAATAAAGTGGGCT 420
QY 118 PheLeuIleArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSer 137
DB 421 TTCATGATCCGAGACATGAGACATGAAGGTTGTTTCTCCCTCTCTGTGCGA----- 474
QY 138 ArgProAlaSerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyr 157
DB 475 -----CACTCAGGGGACACTGTGAACATTACAAATTCGCACACTCGATGATGAGGT 528
QY 158 LeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAlaAspHisTyrSer 177
```

```
DB 529 TTCTTCATTCTACACGGATCCCTTTTCTCTTTGCCAGAGCTGGTACGCCATTATCAA 588
QY 178 GluLeuAlaAspAspIleCysCysLeuLeuLysGluProCys 191
DB 589 GCTAAAGTGGATGGCTTGTGTGTCAGTGCCTTACATACCATGC 630

RESULT 11
US-09-023-655-1080
; Sequence 1080, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1080:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2354 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g182573
; US-09-023-655-1080
```

Alignment Scores:  
Pred. No.: 1.59e-23 Length: 2354  
Score: 312.50 Matches: 75  
Percent Similarity: 52.38% Conservativity: 35  
Best Local Similarity: 35.71% Mismatches: 81  
Query Match: 23.10% Indels: 19  
DB: 4 Gaps: 4

US-09-939-853A-75 (1-261) x US-09-023-655-1080 (1-2354)

```
QY 5 ProSerArgArgLysSerLeuProSerProSerLeuSerSer----- 18
DB 253 CCGTACCCCACTAAGCCCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 312
QY 19 ---SerValGlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAla 37
DB 313 TTCCTCTCTAGGCCATCAACCTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 372
```

QY 38 -----ValAlaLeuGlySerPheProAlaGlyGlyProAlaGlu 50  
 Db 373 GGGATTGGGTGACCCCTGTTCAATGGCCCTGTATCACTATGAGGCTCGAAGTGAAGAC 432  
 QY 51 LeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu---AspGlyAspTrpTrp 69  
 Db 433 CTCACCTTCACCAAGGGCGAAGTCTCCATCTCTGAACAATACTGAAGTGACTGGTGG 492  
 QY 70 ThrValLeuSerGluValSerGlyArgGluTrpAsnIleProSerValHisValGlyLys 89  
 Db 493 GAGCTCGGCTCTCACTCGGCGGAAACCTGGCTGCATCTCCAGCAACTACGTTGGCCCT 552  
 QY 90 Val-----SerHisGlyTrpLeuTrpGluGlyLeuSerArgGluLysAlaGlu 105  
 Db 553 GTTCACTCAATCCAAAGCTGAAGAGTGTCTTTCGAAAGATGGGAGAAAGATGCAGAG 612  
 QY 106 GluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeuArgGluSerGlnThr 125  
 Db 613 AGGCAGCTGCTTTCACAGGCAACCCCGGCGGCTTCTCATTCGGGAAACCGAGACC 672  
 QY 126 ArgArgGlySerTrpSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIle 145  
 Db 673 ACCAAAGTGCTACTCCCTGCTCATCCGGAGTGGATCAGACAGAGCGATCATGTG 732  
 QY 146 ArgHisTrpArgIleHisCysLeuAspAsnGlyTrpLeuTrpIleSerProArgLeuThr 165  
 Db 733 AAGCATTACAGATCCGCAAACTGGACATGGCGGCTACTATACATCACACACGGGTTCAG 792  
 QY 166 PheProSerLeuGlnAlaLeuValAspHisTrpSerGluLeuAlaAspAspIleCysCys 185  
 Db 793 TTCACCTCGGTGAGGAGCTGGTGACCATCATATGAGGTGAATACGGGGCTGTGCAAC 852  
 QY 186 LeuLeuLysGluProCysValLeuGlnArg 195  
 Db 853 CTGCTCATCGCGCCCTGCACCATCATGAAG 882

## RESULT 12

US-09-470-881-2  
 ; Sequence 2, Application US/09470881  
 ; Patent No. 6885938  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CHERESH, David A.  
 ; APPLICANT: ELICERI, Brian  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR MODULATION OF  
 ; TITLE OF INVENTION: ANGIOGENESIS AND VASCULAR PERMEABILITY USING SRC OR  
 ; TITLE OF INVENTION: YES TYROSINE KINASES  
 ; FILE REFERENCE: TSRI 651.2  
 ; CURRENT APPLICATION NUMBER: US/09/470,881  
 ; PRIOR FILING DATE: 1999-12-22  
 ; PRIOR APPLICATION NUMBER: PCT/US99/11780  
 ; PRIOR FILING DATE: 1999-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087,220  
 ; PRIOR FILING DATE: 1998-05-29  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1759  
 ; TYPE: DNA  
 ; ORGANISM: Chicken  
 ; FEATURE:  
 ; NAME/KEY: gene  
 ; LOCATION: (1)..(1759)  
 ; OTHER INFORMATION: chicken c-SRC cDNA  
 ; NAME/KEY: CDS  
 ; LOCATION: (112)..(1710)  
 US-09-470-881-2

## Alignment Scores:

Pred. No.: 3,35e-21 Length: 1759  
 Score: 289.00 Matches: 86  
 Percent Similarity: 47.33% Conservative: 47  
 Best Local Similarity: 30.60% Mismatches: 104  
 Query Match: 21.36% Indels: 45

DB: 4 Gaps: 6  
 US-09-939-853A-75 (1-261) x US-09-470-881-2 (1-1759)  
 QY 5 ProSerArgArgLysSerLeuProSerProSerLeuSerSerSerValGlnGlyGlnGly 24  
 Db 317 CGTTTACCTCGCGGAGCGTGGCGGGGACCTGGCTGGCGGGTCAAC-----363  
 QY 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44  
 Db 364 -----ACTTTCGTGGCTCTCTACGACTACGAG 390  
 QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu 64  
 Db 391 TCCCGGACTGAACCGGACTTGTCTTCAAGAAAGAGAGAACGCTGCAGATTGTCAACAA 450  
 QY 65 ---AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTrpAsnIlePro 83  
 Db 451 ACGGAAGGTGACTGGTGGCTCATTCCTCTACTACAGGACACACGGGCTATACCTCC 510  
 QY 84 SerValHisValGlyLys-----ValSerHisGlyTrpLeuTrpGluGlyLeu 99  
 Db 511 AGTAAGTATGTCGCGCCCTCAGACTCCATCCAGGCTGAAGAGTGTACTTTGGGAAGATC 570  
 QY 100 SerArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeu 119  
 Db 571 ACTCGTCGGAGTCGAGCGGCTGCTCAACCCCGGAAACCCCGGGGAAACCTTCTTG 630  
 QY 120 IleArgGluSerGlnThrArgArgGlySerTrpSerLeuSerValArgLeuSerArgPro 139  
 Db 631 GTCCGGGAGAGCAGACGACAAAGAGTGCCTATTGCTCTCCGTTCTCGACTTTGACAA 690  
 QY 140 AlaSerTrpAspArgIleArgHisTrpArgIleHisCysLeuAspAsnGlyTrpLeuTrp 159  
 Db 691 GCCAAGGGGCTCAATGTGAGACACTACAAGATCCGAGCTGGACAGCGGGCTTCTAC 750  
 QY 160 IleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTrpSerGluLeu 179  
 Db 751 ATCACTCACCACACACAGTTCAGCAGCTGCAGAGCTGGTGGCTACTACTCCAAACAT 810  
 QY 180 AlaAspAspIleCysCysLeuLysGluProCys-----191  
 Db 811 GCTGATGCTTGTGCCCGCGCTGACCAACGCTGCCCCAGCTCAAGCCACGCCAGCCAG 870  
 QY 192 ---ValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrVal 210  
 Db 871 GGAATCGCCAAAGACCGCTGGGAAATCCCGGAGTCCGCTGCGGCTGGAGTGAAGCTG 930  
 QY 211 GlnArg-----ThrProLeuAsn 216  
 Db 931 GGGCAGGCTCTTTGGAGAGTCTGGATGGGGACCTGGAAACGCGCACACAGAGTGGCC 990  
 QY 217 TrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeu 236  
 Db 991 ATAAAGACTCTGAAGCCCGGACCATGTCCCGGAGGCTTCTCTGAGGAAGCCAAAGTG 1050  
 QY 237 LeuSerGluGlyLeuArgGluSerLeu---SerPheTrpIleSerLeuAsnAspGluAla 255  
 Db 1051 ATGAAGAAGTCCCGCATGAGAGCTGTTGACGTGACGAGTGGTGTGGGAAGAGGCC 1110  
 QY 256 Val 256  
 Db 1111 ATC 1113

## RESULT 13

US-07-820-011A-1  
 ; Sequence 1, Application US/07820011A  
 ; Patent No. 5336615  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bell, Leonard  
 ; APPLICANT: Madri, Joseph A.  
 ; APPLICANT: Warren, Stephen L.  
 ; APPLICANT: Luthringer, Daniel J.

TITLE OF INVENTION: Genetically Engineered  
TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced  
TITLE OF INVENTION: Migration  
TITLE OF INVENTION: And Plasminogen Activator Activity  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb storage  
COMPUTER: IBM PC XT  
OPERATING SYSTEM: PC-DOS/MS-DOS 2.10  
SOFTWARE: Displaywrite 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,011A  
FILING DATE: 19920106  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: LB-101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255 1400  
TELEFAX: (203) 254 1101  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1602 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: Double  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE:  
ORGANISM: Gallus, gallus  
PUBLICATION INFORMATION:  
AUTHORS: Takeya, Tatsuo  
AUTHORS: Hanafusa, Hideaburo  
TITLE: Structure and Sequence of the RSV src  
TITLE: Cellular Gene Homologous to the RSV src  
TITLE: Gene and the Mechanism for Generating the  
TITLE: Transforming Virus  
JOURNAL: Cell  
VOLUME: 32  
PAGES: 881-890  
DATE: March, 1983  
US-07-820-011A-1

Alignment Scores:  
Pred. No.: 4.75e-21 Length: 1602  
Score: 287.00 Matches: 81  
Percent Similarity: 51.28% Conservative: 39  
Best Local Similarity: 34.62% Mismatches: 84  
Query Match: 21.21% Indels: 32  
DB: 1 Gaps: 5

US-09-939-853a-75 (1-261) x US-07-820-011A-1 (1-1602)

QY 5 ProSerArgAlaGlySerLeuProSerProSerSerSerValGlnGlyGlnGly 24

DB 206 CCGTTACGTCGCGGACGCTCCCGGGACGCTGGTGGCGGCTACC----- 252

QY 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44

DB 253 -----ACTTTCGTGGCTCTCTACGACTACGAG 279

QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu 64

DB 280 TCCCGAGCTGAACCGGACTTGTCTTCAAGAAAGAGAACCGCTGCAGATTGTCAACAC 339

QY 65 ---AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGlyTrpAsnIlePro 83  
DB 340 ACGGAAGGTGACTGGTGGCTCATTCCTCACTACAGGACAGCGGGCTACATCCCC 399  
QY 84 SerValHisValGlyLys-----ValSerHisGlyTrpLeuTyGluGlyLeu 99  
DB 400 AGTAACATATGTCGGCCCTCAGACTCCATCCAGCTGAAGAGTGGTACTTTTGGGAAGATC 459  
QY 100 SerArgGluLysAlaGluLeuLeuLeuProGlyAsnProGlyGlyValAlaPheLeu 119  
DB 460 ACTCGTCGGGAGTCGAGCGGCTGCTCAACCCGAAACCCCGGGGAACCTTCTTG 519  
QY 120 IleArgGluSerGlnThrArgArgGlySerTySerLeuSerValArgLeuSerArgPro 139  
DB 520 GTCCGGGAGAGCGAGACGACAAAGGTCCTATTCCTCTCCGTTCTCGACTTTGACAAAC 579  
QY 140 AlaSerTrpAspArgIleArgHisTyArgIleHisCysLeuAspAsnGlyTrpLeuTy 159  
DB 580 GCCAAGGGGCTCAATGTGAGCAGCTACAGATCCGCAAGCTGGACAGCGGGCTTCTAC 639  
QY 160 IleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTySerGluLeu 179  
DB 640 ATCACTCAGCAGACAGCTTCAGCAGCTGAGCAGCTGGTGGCTACTACTCAAAACAT 699  
QY 180 AlaAspAspIleCysCysLeuLysGluProCys----- 191  
DB 700 GCTGATGCTTGTGCCCGGCTGACCAACGCTGCCCCACGTCACAGCCCGGACCCAG 759  
QY 192 ---ValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrVal 210  
DB 760 GGACTCGCCAAAGACGCGTGGGAAATCCCGGGAGTGGCTGGCGCTGGAGGTGAAGCTG 819  
QY 211 GlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224  
DB 820 GGGGAGGCTG-CIT---TGGAGAGCTGTGATGGGGACCTG 857

## RESULT 14

PCT-US93-00445-1  
Sequence 1, Application PC/TUS9300445  
GENERAL INFORMATION:  
APPLICANT: Bell, Leonard  
APPLICANT: Madri, Joseph A.  
APPLICANT: Warren, Stephen L.  
APPLICANT: Luthringer, Daniel J.  
TITLE OF INVENTION: Genetically Engineered  
TITLE OF INVENTION: Endothelial Cells  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Maurice M. Klee  
STREET: 1951 Burr Street  
CITY: Fairfield  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06430  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 760 Kb storage  
COMPUTER: DELL 486/50  
OPERATING SYSTEM: DOS 5.0  
SOFTWARE: Displaywrite 3  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00445  
FILING DATE: 19930105  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/820,011  
FILING DATE: 06-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Klee, Maurice M.  
REGISTRATION NUMBER: 30,399  
REFERENCE/DOCKET NUMBER: ALX-101PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 255 1400



ISSUE: 5  
PAGES: 1122-1129  
DATE: May, 1985  
PUBLICATION INFORMATION:  
AUTHORS: Tanaka, Akio  
AUTHORS: Gibbs, Carol P.  
AUTHORS: Arthur, Richard R.  
AUTHORS: Anderson, Stephen K.  
AUTHORS: Kung, Hsing-Jien  
AUTHORS: Fujita, Donald J.  
TITLE: DNA Sequence Encoding the  
TITLE: Amino-Terminal Region of the Human c-src  
TITLE: Protein: Implications of Sequence  
TITLE: Divergence among src-Type Kinase  
TITLE: Oncogenes  
JOURNAL: Molecular and Cellular Biology  
VOLUME: 7  
PAGES: 1978-1983  
ISSUE: 5  
DATE: May, 1987  
US-07-820-011A-3

Alignment Scores:  
Pred. No.: 5.41e-20 Length: 1611  
Score: 276.50 Matches: 87  
Percent Similarity: 46.15% Conservative: 45  
Best Local Similarity: 30.42% Mismatches: 117  
Query Match: 20.44% Indels: 38  
DB: 1 Gaps: 6

US-09-939-853a-75 (1-261) x US-07-820-011A-3 (1-1611)

QY 8 ArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGlnGlyProValThr 27  
DB 155 CGGCTTCGCCCCCGCGCGCCGAGCCCAAGCTGTCGGAGGCTTCACTCCTCGGACA 214  
QY 28 MetGluAlaGluArgSerLysAla-----ThrAlaValAla 39  
DB 215 CGTCACCTCCCGCGAG-AGGCGCGCGCGCGTGGCGGTGAGTGACCACTTTGTGGCC 273  
QY 40 LeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeu 59  
DB 274 CTCATGACTATGACTAGACGGAGACACACTGTCTTCAAGAAGGCGAGCGCTC 333  
QY 60 ThrIleValSerGlu---AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArg 78  
DB 334 CAGATTGTCAACACACACAGAGGAGACTGTGTGGTGGCCACTCGCTCAGCAGAGGACAG 393  
QY 79 GluTyrAsnIleProSerValHisValGlyLys-----ValSerHisGlyTrp 94  
DB 394 ACAGCTACATCCCGAGCACTAGCTGGCGGCTCCGACTCCATCCAGGCTGAGGAGTGG 453  
QY 95 LeuTyrGluGlyLeuSerArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnPro 114  
DB 454 TATTTGGCAAGATCACCACAGCGGAGTCAGAGCGGTACTGTCTCAATGCAGAGAACCCG 513  
QY 115 GlyGlyAlaPheLeuIleArgGluSerGlnThrArgGlySerTyrSerLeuSerVal 134  
DB 514 AGAGGGACCTTCCTCGTCGGAAGAAGTAGACCAACGAAAGGTGCTACTGCTCTCAGTG 573  
QY 135 ArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyrArgIleHisCysLeuAsp 154  
DB 574 TCTGACTTCGACAAAGCGGCTCAACGCTGAGCAGCTACAGATCCGCAAGCTGGAC 633  
QY 155 AsnGlyTrpLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAsp 174  
DB 634 AGCGCGGCTTCTACATCACCCTCCGACCCAGGTTCACAGCCCTGCAGCAGCTGGTGGCC 693  
QY 175 HisTyrSerGluLeuAlaAspIleCysCysLeuLeuLysGluProCys----- 191  
DB 694 TACTACTCAACACAGCGGAGTGGCTGTGCCACCGCTCACCACCGGTGTGCCACCGTCC 753  
QY 192 -----ValLeuGlnArgAlaGlyProLeuProGlyLysAspIlePro 205

Search completed: February 20, 2004, 11:33:00  
Job time : 120 secs

DB 754 AAGCCGCACTCAGGGCTCGCCCAAGGATCCCTGGGAGATCCCTCGGAGTGGCTCGG 813  
QY  
DB 206 LeuProValThrValGlnArg-----  
DB 814 CTGGAGGTCAAGCTGGCGCAGGCTGCTTTGGCGAGGTGTGATGGGACCTTGGAAACGGT 873  
QY 213 ----ThrProLeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThr 231  
DB 874 ACCACCGAGGTGGCCATCAAAACCCCTGAAGCTGGCACGATGTCTCCAGAGGCTTCTCG 933  
QY 232 GlyGluGluSerLeuLeuSerGluGlyLeuArgGluSerLeu---SerPheTyrIleSer 250  
DB 934 CAGGAGGCCAGGTCTATGAGAGAGCTGAGGATGAGAAAGCTGGTGCACATTGTATGCTGTG 993  
QY 251 LeuAsnAspGluAlaVal 256  
DB 994 GTTTCAGAGGAGCCCAATT 1011

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2004, 10:48:39 ; Search time 357 Seconds

(without alignments)  
2560.026 Million cell updates/sec

Title: US-09-939-853A-75

Perfect score: 1353

Sequence: 1 MGSUFRKSLPSLSSVV.....RESLSFYISLNDRAVSLDDA 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delext 7.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO spool\_p/US09939853/runat\_19022004\_145341\_24590/app\_query.fasta\_1.455  
-DB=Published Applications\_NA\_QMWT-fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSZE=500 -MINLEN=0  
-MAXLEN=2000000000 -MMAP=US09939853 @CGN 1 1 333 @runat\_19022004\_145341\_24590  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/prodata/2/pubpna/PTC\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/prodata/2/pubpna/US05\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/prodata/2/pubpna/PTCUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/prodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/prodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/prodata/2/pubpna/US09D\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/prodata/2/pubpna/US10D\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	1347	99.6	786	14	US-10-043-649-1	Sequence 1, Appli
2	826	61.0	783	9	US-09-867-550-953	Sequence 953, App
3	643	47.5	864	10	US-09-814-353-21302	Sequence 21302, A
4	586	43.3	875	9	US-09-867-550-1915	Sequence 1915, App
5	488	36.1	3756	13	US-10-002-600-91	Sequence 91, Appl
6	487	36.0	2655	9	US-09-954-456-499	Sequence 499, App
7	452.5	33.4	444	9	US-09-867-550-951	Sequence 951, App
8	370.5	27.4	2298	14	US-10-175-523-50	Sequence 50, Appl
9	370.5	27.4	2298	15	US-10-159-563-343	Sequence 343, App
10	360.5	26.6	1924	15	US-10-193-720-1	Sequence 1, Appli
11	360.5	26.6	2015	9	US-09-954-456-1993	Sequence 1993, Ap
12	360.5	26.6	2015	14	US-10-007-010-3	Sequence 3, Appli
13	360.5	26.6	2341	14	US-10-252-157-140	Sequence 140, App
14	360.5	26.6	2343	15	US-10-062-674-2038	Sequence 2038, Ap
15	358.5	26.5	1911	9	US-09-917-8008-1611	Sequence 1611, Ap
16	350	25.9	320	10	US-09-814-353-17314	Sequence 17314, A
17	348	25.7	2032	15	US-10-366-288-27	Sequence 27, Appl
18	340	25.1	2017	15	US-10-062-674-1776	Sequence 1776, Ap
19	340	25.1	2129	10	US-09-960-706-954	Sequence 954, App
20	340	25.1	2129	15	US-10-305-720-1452	Sequence 1452, Ap
21	324	23.9	2433	14	US-10-240-965-114	Sequence 114, App
22	322.5	23.8	2451	9	US-09-771-161A-4	Sequence 4, Appli
23	320	23.7	1609	9	US-09-771-161A-30	Sequence 30, Appl
24	320	23.7	1995	9	US-09-771-161A-31	Sequence 31, Appl
25	320	23.7	4414	14	US-10-101-510-512	Sequence 512, App
26	315.5	23.3	4343	15	US-10-159-563-184	Sequence 184, App
27	315.5	23.3	4517	14	US-10-298-377A-3	Sequence 3, Appli
28	315.5	23.3	4517	14	US-10-007-926A-304	Sequence 304, App
29	314.5	23.2	549	10	US-09-918-995-25044	Sequence 25044, A
30	312.5	23.1	1590	15	US-10-085-117-18	Sequence 18, Appl
31	312.5	23.1	2354	9	US-09-967-768A-300	Sequence 300, App
32	312.5	23.1	2354	15	US-10-353-690-123	Sequence 123, App
33	312.5	23.1	2354	15	US-10-085-117-17	Sequence 17, Appl
34	303	22.4	1554	15	US-10-085-117-15	Sequence 15, Appl
35	303	22.4	2179	15	US-10-085-117-14	Sequence 14, Appl
36	293.5	21.0	488	10	US-09-918-995-19745	Sequence 19745, A
37	246	18.2	432	9	US-09-864-761-2829	Sequence 2829, Ap
38	246	18.2	448	9	US-09-864-761-15513	Sequence 15513, A
39	240.5	17.8	2770	9	US-09-977-269-5	Sequence 5, Appli
40	240.5	17.8	2770	10	US-09-977-261-5	Sequence 5, Appli
41	240.5	17.8	2770	10	US-09-977-261-5	Sequence 5, Appli
42	240.5	17.8	2863	9	US-09-954-456-1631	Sequence 1631, Ap
43	240.5	17.8	7607	9	US-09-982-610-19	Sequence 19, Appl
44	238.5	17.6	2173	15	US-10-094-749-1087	Sequence 1087, Ap
45	230.5	17.0	342	15	US-10-062-674-1242	Sequence 1242, Ap

## ALIGNMENTS

### RESULT 1

US-10-043-649-1  
; Sequence 1, Application US/10043649  
; Publication No. US20030059924A1  
; GENERAL INFORMATION:  
; APPLICANT: Holland, Sacha J.  
; APPLICANT: Menderhall, Marcy K.  
; APPLICANT: Pardo, Jorge  
; APPLICANT: Spencer, Collin  
; APPLICANT: Fu, C. Alan  
; APPLICANT: Luo, Ying  
; APPLICANT: Payan, Donald G.  
; APPLICANT: Mancebo, Helena S.Y.  
; APPLICANT: Wu, Jun  
; APPLICANT: Zhou, Xiulan  
; APPLICANT: Shen, Mary  
; APPLICANT: Liao, X. Charlene  
; APPLICANT: Sheng, Ning  
; TITLE OF INVENTION: Cloning of a No. US20030059924A1e1 Inhibitor of Antigen-receptor  
; FILE REFERENCE: A-70219-1/RMS/DHR  
; CURRENT APPLICATION NUMBER: US/10/043,649  
; CURRENT FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/260,953

; PRIOR FILING DATE: 2001-01-10  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 786  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(786)  
 ; OTHER INFORMATION:  
 US-10-043-649-1

Alignment Scores:  
 Pred. No.: 63le-146 Length: 786  
 Score: 1347.00 Matches: 260  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.56% Indels: 0  
 DB: 14 Gaps: 0

US-09-939-853A-75 (1-261) x US-10-043-649-1 (1-786)

Qy	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db	1	ATGGGAAGTCTGCCAGCAGAGAGAAATCTCTCCCAAGCCCAAGCTTGAGTTCTCTCTGTC	60
Qy	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	61	CNAGGCCAGGAGCTGTGACCATGTGAAGCAGAGAGAGAGAGCCACAGCCGTGGCCCTG	120
Qy	41	GlySerPheProAlaGlyProAlaGluLeuSerLeuArgLeuGluProLeuThr	60
Db	121	GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTCTGAGACTCGGGAGCCATTGACC	180
Qy	61	IleValSerGluAspGlyAspTrrThrValLeuSerGluValSerGlyArgGluTyr	80
Db	181	ATGCTCTCTGAGATGAGACTGTGTGACCGTGTCTGTAAGTCTCAGCGAGAGATAT	240
Qy	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
Db	241	AACATCCCAAGCTCCACGTGGCCCAAGTCTCCATGGTGGTGTATGAGGCCCTGAGC	300
Qy	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyAlaPheLeuLeu	120
Db	301	AGGAGAAAGCAGAGAACTGCTGTGTACCTGGGAACCTTGAGGGCCCTTCTCTATC	360
Qy	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	361	CGGAGAGCCAGACAGAGAGCTCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGCA	420
Qy	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIle	160
Db	421	TCCTGGACCCGATCAGACTACAGATCCACTGCTTTGACAATGGCTGGCTGTATATC	480
Qy	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
Db	481	TCACCGGCTCTACCTTCCCTTACTCCAGCCCTGGTGGACCATTTACTCTGAGCTGGC	540
Qy	181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
Db	541	GATGACATCTGCTGCTTACTCAAGGAGCCCTGTCTCTGACAGAGGCTGGCCGCTCCCT	600
Qy	201	GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu	220
Db	601	GGCAGAGATATACCTTACTCTGAGCTGTGTCAGAGAGACACCATCACTCACTGGAAGAGCTG	660
Qy	221	AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuSerGluGly	240
Db	661	GACAGCTCCCTCTGTTCTTCTGAAGCTGCCACAGGGAGAGGCTCTTCTCAGTGAGGT	720
Qy	241	LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp	260

Db	721	CTCCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGGCTGTCTCTTTTGGATGAT	780
Qy	261	Ala 261	
Db	781	GCC 783	

RESULT 2  
 US-09-867-550-953  
 ; Sequence 953, Application US/09867550  
 ; Patent No. US20020082206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad,  
 ; APPLICANT: Conley, Pamela,  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
 ; FILE REFERENCE: 21402-013 (Cura-313)  
 ; CURRENT APPLICATION NUMBER: US/09/867,550  
 ; CURRENT FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: USN 60/208,427  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 953  
 ; LENGTH: 763  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-867-550-953

Alignment Scores:  
 Pred. No.: 9.35e-86 Length: 763  
 Score: 826.00 Matches: 158  
 Percent Similarity: 99.37% Conservative: 0  
 Best Local Similarity: 99.37% Mismatches: 1  
 Query Match: 61.05% Indels: 0  
 DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x US-09-867-550-953 (1-763)

Qy	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
Db	286	ATGGGAAGTCTGCCAGCAGAGAAATCTCTCCCAAGCCCAAGCTTGAGTTCTCTCTGTC	345
Qy	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
Db	346	CAAGGCCAGGAGCTGTGACCATGTGAAGCAGAGAGAGAGAGCCACAGCCGTGGCCCTG	405
Qy	41	GlySerPheProAlaGlyProAlaGluLeuSerLeuArgLeuGluProLeuThr	60
Db	406	GGCAGTTTCCCGCAGGTGGCCCGCGAGCTGTCTGAGACTCGGGAGCCATTGACC	465
Qy	61	IleValSerGluAspGlyAspTrrThrValLeuSerGluValSerGlyArgGluTyr	80
Db	466	ATCGTCTCTGAGGATGAGACTGTGTGACCGTGTCTGTAAGTCTCAGCGAGAGATAT	525
Qy	81	AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer	100
Db	526	AACATCCCAAGCTCCACGTGGCCCAAGTCTCCATGGTGGTGTATGAGGGCTCAGC	585
Qy	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu	120
Db	586	AGGAGAAAGCAGAGAACTGCTGTGTACTTACCTGGGAGCCCTTGAGGGGCTTCTCTATC	645
Qy	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
Db	646	CGGAGAGCCAGACAGAGAGAGGCTCTTACTCTGTCTGTCGCGCTCAGCCGCCCTGCA	705
Qy	141	SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyr	159
Db	706	TCCTGGGACCGATCAGACTACAGATCCACTGCTTGCATTCAGTGGCTGGTGTAC	762

RESULT 3

US-09-814-353-21302  
 ; Sequence 21302, Application US/09814353  
 ; Publication No. US20030165831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lee, John  
 ; APPLICANT: Thompson, Pamela  
 ; APPLICANT: Lillie, James  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
 ; FILE REFERENCE: MRI-006B  
 ; CURRENT APPLICATION NUMBER: US/09/814,353  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/191,031  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/207,124  
 ; PRIOR FILING DATE: 2000-05-25  
 ; PRIOR APPLICATION NUMBER: US 60/211,940  
 ; PRIOR FILING DATE: 2000-06-15  
 ; PRIOR APPLICATION NUMBER: US 60/216,820  
 ; PRIOR FILING DATE: 2000-07-07  
 ; PRIOR APPLICATION NUMBER: US 60/220,661  
 ; PRIOR FILING DATE: 2000-07-25  
 ; PRIOR APPLICATION NUMBER: US 60/257,672  
 ; PRIOR FILING DATE: 2000-12-21  
 ; NUMBER OF SEQ ID NOS: 22037  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 21302  
 ; LENGTH: 864  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: 1, 2, 3, 32, 862, 863, 864  
 ; OTHER INFORMATION: n = A,T,C or G

Alignment Scores:

Pred. No.: 1,54e-64 Length: 864  
 Score: 643.00 Matches: 126  
 Percent Similarity: 98.45%  
 Best Local Similarity: 97.67%  
 Query Match: 47.52%  
 DB: 10 Gaps: 0

US-09-939-853A-75 (1-261) x US-09-814-353-21302 (1-864)

QY 1 MetGlySerLeuProSerArgA:GlySerLeuProSerProSerLeuSerSerVal 20  
 Db 450 ATGGGAAGTCTGCCAGCAGAGAAATCTCTGCCAAGCCAGCTTGATTCCTCTGTC 509  
 QY 21 GluGlyGlnGlyProValThrMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40  
 Db 510 CAAGGCCAGGAGCTGTGACCATGGAAGCAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 569  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 Db 570 GGCAAGTTCGGCAGAGTGGCCGGCGGAGCTGTCTGAGACTCGGGAGCCATTGACC 629  
 QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 Db 630 ATCGTCTCTGAGGATGAGACTGGTGGACGGTCTGTCTGAAGTCTCAGGCAGAGAGTAT 689  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuThrGluGlyLeuSer 100  
 Db 690 AACATCCCCAGGCTCCACGGTGGCAAGTCTCCATGGGTGGCTGTATGAGGCCCTGAGC 749  
 QY 101 ArgGluLysAlaGluLeuLeuLeuLeuLeuLeuProGlyAsnProGlyAlaPheLeuIle 120  
 Db 750 AGGGAGAAACAGAGAACTGCTGTTTACCTGGGAAACCTGGAGGGGCCCTTCTCATC 809  
 QY 121 ArgGluSerGlnThrArgArgGlySer 129

RESULT 4

Db 810 CGGAGACCCAGACCAAGAGAGTCC 836  
 ; Sequence 1915, Application US/09867550  
 ; Patent No. US20020082206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad,  
 ; APPLICANT: Conley, Pamela  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and i  
 ; TITLE OF INVENTION: Thereby  
 ; FILE REFERENCE: 21402-013 (Cura-313)  
 ; CURRENT APPLICATION NUMBER: US/09/867,550  
 ; CURRENT FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: USSN 60/208,427  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1915  
 ; LENGTH: 875  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Wherein n is one of a or t or c or g

Alignment Scores:

Pred. No.: 6,02e-58 Length: 875  
 Score: 586.00 Matches: 112  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 43.31%  
 DB: 9 Gaps: 0

US-09-939-853A-75 (1-261) x US-09-867-550-1915 (1-875)

QY 150 IleHisCysLeuAspAsnGlyTrpLeuTyIleSerProArgLeuThrPheProSerLeu 169  
 Db 4 ATCCACTGCTTGACAAATGGCTGTGATCATCTCACCGCGCCTCACCTTCCCTCACTC 63  
 QY 170 GlnAlaLeuValAspHisTyrSerGluLeuAlaAspIleCysCysLeuLeuLysGlu 189  
 Db 64 CAGGCCCTGGTGGACCATTAATCTCTGAGCTGGCGGATGACATCTGCTGCTACTCAAGGAG 123  
 QY 190 ProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThr 209  
 Db 124 CCCTGTGTCTGCGAGAGGGCTGGCCCGCTCCCTGGCAAGGATATACCCCTACCTGTGACT 183  
 QY 210 ValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeuLeuPheSerGluAla 229  
 Db 184 CTGCAGAGACACACATCACTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAGACT 243  
 QY 230 AlaThrGlyGluGluSerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIle 249  
 Db 244 GCCACAGGGAGGAGTCTCTTCTCAGTGAGGGTCTCCGGAGTCCCTCAGCTTCTATCATC 303  
 QY 250 SerLeuAsnAspGluAlaValSerLeuAspAla 261  
 Db 304 AGCCTGATGACGAGGGTGTCTCTTTGGATGATGCC 339

RESULT 5

US-10-002-600-91  
 ; Sequence 91, Application US/10002600  
 ; Publication No. US20020137077A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hopkins, Christopher M.  
 ; APPLICANT: Peterson, David P.

; APPLICANT: Cocks, Benjamin G.  
 ; APPLICANT: Hawkins, Phillip R.  
 ; TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS  
 ; FILE REFERENCE: PA-0042 US  
 ; CURRENT APPLICATION NUMBER: US/10/002,600  
 ; CURRENT FILING DATE: 2001-10-25  
 ; PRIOR APPLICATION NUMBER: 60/243,521  
 ; PRIOR FILING DATE: 2000-10-25  
 ; NUMBER OF SEQ ID NOS: 116  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 91  
 ; LENGTH: 3756  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Template ID: 059263.15  
 US-10-002-600-91

Alignment Scores:  
 Pred. No.: 9,54e-46 Length: 3756  
 Score: 488.00 Matches: 101  
 Percent Similarity: 57.20% Conservative: 46  
 Best Local Similarity: 39.30% Mismatches: 94  
 Query Match: 36.07% Indels: 16  
 DB: 13 Gaps: 3

US-09-939-853A-75 (1-261) x US-10-002-600-91 (1-3756)  
 QY 5 ProSerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24  
 DB 1098 CCAAGGAAAGAAAGAAATGGGAACAGCATGAAATCCACCCCTCGCTCCGAGAGG 1157  
 QY 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44  
 DB 1158 CCCTGCGCCACCCGAGGAGCTGGATAGCGACTTCTCCCGTGTAAAGTACACCG 1217  
 QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrLeuValSerGlu 64  
 DB 1218 TCTCTGACATCAGCCCGCGATATTCGCGGAGGAGAACTGCGTGTGATTTCTGAT 1277  
 QY 65 AspGlyAspTrpThrValLeuSerGluValSerGlyValArgGlyLysLeuProSer 84  
 DB 1278 GAAGGGGCTGGTGAAGAGTATTTCTTAGCAGTGTTCAGAGAGTATCATCTCTGA 1337  
 QY 85 ValHisValGlyLysValSerHisGlyTrpLeuTrpGlyLeuSerArgGlyLysAla 104  
 DB 1338 ATATGTGCGCCAGAGTTTACTCATGCTGCTGTTTCCGAGGAGTATCATCTCTGA 1397  
 QY 105 GluGluLeuLeuLeuLeuProGlyAsnProGlyAlaPheLeuLeuArgGlySerGln 124  
 DB 1398 GAGAGCTGCTGAGCTGCCAGACACAAAGTGGCTTCTTATGATCAGAGAGTGTAG 1457  
 QY 125 ThrArgArgGlySerTrpSerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144  
 DB 1458 ACCAAGAAGGTTTACTCATCTGCTGCTGAGACAAAGG-----CAG 1499  
 QY 145 IleArgHisTrpArgIleHisCysLeuAspAsnGlyTrpLeuTrpLeuSerProArgLeu 164  
 DB 1500 GTAAGCATATTACCGCATTTTCGCTGCCCAACACTGGTACTACATATTTCCCGAGGCTC 1559  
 QY 165 ThrPheProSerLeuGlnAlaLeuValAlaPheHisTrpSerGluLeuAlaAspAspIleCys 184  
 DB 1560 ACCTTCAGTCCCTGGAGGACTTGTGAACCACTATTCTGAGTGGCTGATGCCCTGTGC 1619  
 QY 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
 DB 1620 TGTGTGCTCACACGCGCTGCTGACACAAAGCAGCGCTGCCCGAGCAGTGTGAGGCTCC 1679  
 QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224  
 DB 1680 AGCTCACCTGTTCACCTTGCCTCAGAGACTGTGAGTGGAGAGTGTCTCAGA----- 1733

QY 225 LeuPheSerGluAlaAlaThrGly-----GluGlySerLeuLeu 237  
 DB 1734 --CTGAGGAGGAGCCCGAGGAGACAGAGAACCCCTTGGGTGAGCAGGAGTCCCTTTTC 1790  
 QY 238 SerGlyGlyLeuArgGlySerLeuSerPheTrpLeuSerLeuAsnAspGlu 254  
 DB 1791 AGCTATGGCTTCGAGAGAGCATTCCTTACTGCTCCCTGACCATGAG 1841

RESULT 6  
 US-09-954-456-499  
 ; Sequence 499, Application US/09954456  
 ; Patent No. US20020115057A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
 ; FILE REFERENCE: 689290-76  
 ; CURRENT APPLICATION NUMBER: US/09/954,456  
 ; CURRENT FILING DATE: 2001-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/233,617  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/234,052  
 ; PRIOR FILING DATE: 2000-09-20  
 ; PRIOR APPLICATION NUMBER: US/60/234,923  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235,134  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235,637  
 ; PRIOR FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US/60/235,638  
 ; PRIOR FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US/60/235,711  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,720  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,840  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,863  
 ; PRIOR FILING DATE: 2000-09-27  
 ; NUMBER OF SEQ ID NOS: 2276  
 ; SOFTWARE: Patent in version 3.0  
 ; SEQ ID NO 499  
 ; LENGTH: 2665  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-954-456-499

Alignment Scores:  
 Pred. No.: 7,73e-46 Length: 2665  
 Score: 487.00 Matches: 101  
 Percent Similarity: 57.20% Conservative: 46  
 Best Local Similarity: 39.30% Mismatches: 94  
 Query Match: 35.99% Indels: 16  
 DB: 9 Gaps: 3

US-09-939-853A-75 (1-261) x US-09-954-456-499 (1-2665)  
 QY 5 ProSerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyGlnGly 24  
 DB 24 CAGGGAAGAAAGAAATGGGAACAGCATGAAATCCACCCCTCGCTCCGAGAGG 83  
 QY 25 ProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPhePro 44  
 DB 84 CCCTGCGCCACCCGAGGAGCTGGATAGCGACTTCTTGGCTGCTAAGTACTACCG 143  
 QY 45 AlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGlu 64  
 DB 144 TCTCTGACATCAGCAGCCCGCGATATTCGCGAGGAGGAGAACTGCGTGTGATTTCTGAT 203  
 QY 65 AspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGlyLysTrpAsnIleProSer 84  
 DB 204 GAAGGGGCTGTGTGAAAGCTATTCTTCTAGCAGTGTGTCGAGAGAGTTACATCCCTGGA 263

QY 85 ValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSerArgGluLysAla 104  
 ...  
 Db 264 ATATGTGTGGCCAGAGATTACCATGGCTGGCTGTGGGGCTGGGAGAGACAGGCC 323  
 ...  
 QY 105 GluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeuLeuLeuGln 124  
 ...  
 Db 324 GAGGAGCTGCTGACCTGCCAGACACAAAGGTGGCTCTTCATGATCAGAGAGATGAG 383  
 ...  
 QY 125 ThrArgArgGlySerTySerLeuSerValArgLeuSerArgProAlaSerTrpAspArg 144  
 ...  
 Db 384 ACCAAGAAGGGTTTACTCACTCGGTGAGA-----CACAGGCAG 425  
 ...  
 QY 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrIleSerProArgLeu 164  
 ...  
 Db 426 GAAAGCATACCGATTTCCTGTCGCGAACAACCTGTACTACATTTCCCGAGGCTC 485  
 ...  
 QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCys 184  
 ...  
 Db 486 ACCTCCAGTGGCTGGAGGACCTGGTGAACCACTATTCTGAGTGGCTGATGGCTGTGC 545  
 ...  
 QY 185 CysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204  
 ...  
 Db 546 TGTGTGCTCACCAGCCCTGCTCAGAAAGCTGTGGACTGGAGGAGAGTGTCCAGA----- 659  
 ...  
 QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeuAspSerSerLeu 224  
 ...  
 Db 606 AGCTACCTGTCACTTGGCTCAGAAAGCTGTGGACTGGAGGAGAGTGTCCAGA----- 659  
 ...  
 QY 225 LeuPheSerGluAlaAlaThrGly-----GluGluSerLeuLeu 237  
 ...  
 Db 660 ---CTGCAGAGAGACCCCGAGGAGAACAGAGAACCCGCTGGGGTAGAGAGTCCCTTTTC 716  
 ...  
 QY 238 SerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu 254  
 ...  
 Db 717 AGCTATGGCTTCGAGAGAGCATTGCTCTTACCTGTCCCTGACCAAGTGTGAG 767  
 ...  
 RESULT 7  
 US-09-867-550-951  
 ; Sequence 951, Application US/09867550  
 ; Patent No. US20020082206A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leach, Martin D.  
 ; APPLICANT: Mehraban, Fuad,  
 ; APPLICANT: Conley, Pamela  
 ; APPLICANT: Law, Debbie  
 ; APPLICANT: Topper, James  
 ; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
 ; FILE REFERENCE: 21402-013 (Cura-313)  
 ; CURRENT APPLICATION NUMBER: US/09/867,550  
 ; PRIOR FILING DATE: 2001-09-20  
 ; PRIOR APPLICATION NUMBER: USSN 60/208,427  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 2125  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 951  
 ; LENGTH: 444  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-867-550-951  
 Alignment Scores:  
 Pred. No.: 6,21e-43 Length: 444  
 Score: 452.50 Matches: 96  
 Percent Similarity: 76.80% Conservative: 0  
 Best Local Similarity: 76.80% Mismatches: 29  
 Query Match: 33.44% Indels: 9  
 DB: 1 Gaps: 1  
 US-09-939-853A-75 (1-261) x US-09-867-550-951 (1-444)  
 QY 1 MetGlySerLeuProSerArgGlySerLeuProSerProSerLeuSerSerVal 20  
 ...

Db 157 ATCGGAAGTCTGCCAGCAGACAGAAATCTTCCAGAGCCCAAGCTTGAGTTCTCTCTC 216  
 ...  
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40  
 ...  
 Db 217 CAAGGCAGGACCTGTGACCATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 276  
 ...  
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60  
 ...  
 Db 277 GGAGATTTCCTGGCAGGTGGCCCGCCGAGCTGTCTGAGACTCGGAGAGCATTGACC 336  
 ...  
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80  
 ...  
 Db 337 ATCGTCTCTGAG----- 348  
 ...  
 QY 81 AsnIleProSerValHisValGlyLysValSerHisGlyTrpLeuTyrGluGlyLeuSer 100  
 ...  
 Db 349 -----TGGCTGTATGAGGCGCTGAGC 369  
 ...  
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuLeu 120  
 ...  
 Db 370 AGGAGAAACAGAGAACTGTCTGTGTGTACCTGGGAACCTCGAGGGGCTTCTCTCATC 429  
 ...  
 QY 121 ArgGluSerGlnThr 125  
 ...  
 Db 430 CGGAGAGCCAGACC 444  
 ...  
 RESULT 8  
 US-10-175-523-50  
 ; Sequence 50, Application US/10175523  
 ; Publication No. US20030096264A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brockman, Jeffrey  
 ; APPLICANT: Evans, David  
 ; APPLICANT: Hook, Derek  
 ; APPLICANT: Klimczak, Leszek  
 ; APPLICANT: Laeng, Pascal  
 ; APPLICANT: Palfreyman, Michael  
 ; APPLICANT: Rajan, Prithi  
 ; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)  
 ; FILE REFERENCE: 3235/1J795-US3  
 ; CURRENT APPLICATION NUMBER: US/10/175,523  
 ; CURRENT FILING DATE: 2002-06-18  
 ; PRIOR APPLICATION NUMBER: US 60/299,151  
 ; PRIOR FILING DATE: 2001-06-18  
 ; PRIOR APPLICATION NUMBER: US 60/317,828  
 ; PRIOR FILING DATE: 2001-09-07  
 ; PRIOR APPLICATION NUMBER: US 60/325,150  
 ; PRIOR FILING DATE: 2001-09-25  
 ; PRIOR APPLICATION NUMBER: US 60/333,047  
 ; PRIOR FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: US 60/349,936  
 ; PRIOR FILING DATE: 2002-01-18  
 ; PRIOR APPLICATION NUMBER: US 60/361,834  
 ; PRIOR FILING DATE: 2002-03-04  
 ; NUMBER OF SEQ ID NOS: 197  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 50  
 ; LENGTH: 2298  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-175-523-50  
 Alignment Scores:  
 Pred. No.: 1.81e-32 Length: 2298  
 Score: 370.50 Matches: 80  
 Percent Similarity: 57.71% Conservative: 36  
 Best Local Similarity: 59.60% Mismatches: 76  
 Query Match: 27.38% Indels: 9  
 DB: 14 Gaps: 3  
 US-09-939-853A-75 (1-261) x US-10-175-523-50 (1-2298)  
 QY 6 SerArgGlySerLeuProSerProSerLeuSerSerValGlnGlyPro 25  
 ...

```
Db 409 TCCAAATAACAGCAAGCCAGTCCAGAA---TCTCAGCTTTTACCTGGACAGAGGTTT 465
Qy 26 ValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPheProAla 45
Db 466 CAACCTAAAGATCCAGAGGAACAGGAGACATTTGGTAGCCTTGACCCCTATGATGGC 525
Qy 46 GlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGluAsp 65
Db 526 ATCCACCGGACGACTTGTCTTTCAAGAAAGGAGAGAGATGAAGTCCCTGGAGGAGCAT 585
Qy 66 GlyAspTrpThrValLeuSerGluValSerGlyArgGluTrpAsnIleProSerVal 85
Db 586 GGAGAATGGTGGAAAGCAAGATCCCTTTTACAAAAAGAGAGGCTTCATCCCCAGCAAC 645
Qy 86 HisValGlyLysVal-----SerHisGlyTrpLeuTrpGluGlyLeuSerArg 101
Db 646 TATGTGCCAAACTCAACACTTAGAAACAGAGAGTGGTTTTTCAAGGATATAACACAG 705
Qy 102 GluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArg 121
Db 706 AAGACCGCAAGAGCGAGCTTTTGGCCAGGAAATAGCGCTGGAGCTTTCTCTATTAGA 765
Qy 122 GluSerGlnThrArgArgGlySerTySerLeuSerValArgLeuSerArgProAlaSer 141
Db 766 GAAGTGAACATTAAAGAGAGCTTCTCTCTGTCTGTCTGAGACTTTGACCTGTGCAT 825
Qy 142 TrpAspArgIleArgHisTyArgIleHisCysLeuAspAsnGlyTrpLeuTyIleSer 161
Db 826 GGTGATGTTATTAGCACTACAAATAGAAAGTCTGGATAATGGGGCTATTACATCTCT 885
Qy 162 ProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTySerGluLeuAlaAsp 181
Db 886 CCAGATACATTTTCCCTGTATCAGCGACATGATTAACATTACCAAGAGAGGAGAT 945
Qy 182 AspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGly 201
Db 946 GGCTTGTGCAGAAAGATTGGAGAGGCTTGTTATT-----AGTCCCAAGCCACAG 993

RESULT 9
US-10-159-563-343
; Sequence 343, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 343
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-343

Alignment Scores:
Pred. No.: 1,818-32 Length: 2298
Score: 370.50 Matches: 80
Percent Similarity: 57.71% Conservative: 36
Best Local Similarity: 39.80% Mismatches: 76
Query Match: 27.38% Indels: 9
```

```
DB: 15 Gaps: 3
US-09-939-853A-75 (1-261) x US-10-159-563-343 (1-2298)
Qy 6 SerArgArgLysSerLeuProSerProSerLeuSerSerValGlnGlyGlnGlyPro 25
Db 409 TCCAAATAACAGCAAGCCAGTCCAGAA---TCTCAGCTTTTACCTGGACAGAGGTTT 465
Qy 26 ValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeuGlySerPheProAla 45
Db 466 CAACCTAAAGATCCAGAGGAACAGGAGACATTTGGTAGCCTTGATCCCTATGATGGC 525
Qy 46 GlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThrIleValSerGluAsp 65
Db 526 ATCCACCGGACGACTTGTCTTTCAAGAAAGGAGAGAGATGAAGTCCCTGGAGGAGCAT 585
Qy 66 GlyAspTrpThrValLeuSerGluValSerGlyArgGluTrpAsnIleProSerVal 85
Db 586 GGAGAATGGTGGAAAGCAAGTCCCTTTTACAAAAAGAGAGGCTTCATCCCCAGCAAC 645
Qy 86 HisValGlyLysVal-----SerHisGlyTrpLeuTrpGluGlyLeuSerArg 101
Db 646 TATGTGCCAAACTCAACACTTAGAAACAGAGAGTGGTTTTTCAAGGATATAACACAG 705
Qy 102 GluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIleArg 121
Db 706 AAGACCGCAAGAGCGAGCTTTTGGCCAGGAAATAGCGCTGGAGCTTTCTCTATTAGA 765
Qy 122 GluSerGlnThrArgArgGlySerTySerLeuSerValArgLeuSerArgProAlaSer 141
Db 766 GAAGTGAACATTAAAGAGAGCTTCTCTCTGTCTGTCTGAGACTTTGACCTGTGCAT 825
Qy 142 TrpAspArgIleArgHisTyArgIleHisCysLeuAspAsnGlyTrpLeuTyIleSer 161
Db 826 GGTGATGTTATTAGCACTACAAATAGAAAGTCTGGATAATGGGGCTATTACATCTCT 885
Qy 162 ProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTySerGluLeuAlaAsp 181
Db 886 CCAGATACATTTTCCCTGTATCAGCGACATGATTAACATTACCAAGAGAGGAGAT 945
Qy 182 AspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuProGly 201
Db 946 GGCTTGTGCAGAAAGATTGGAGAGGCTTGTTATT-----AGTCCCAAGCCACAG 993

202 Lys 202
994 AAG 996

RESULT 10
US-10-193-720-1
; Sequence 1, Application US/10193720
; Publication No. US20040009173A1
; GENERAL INFORMATION:
; APPLICANT: Frieria, Annabelle M.
; APPLICANT: Wong, Brian R.
; APPLICANT: Masuda, Etsaban
; APPLICANT: Powell, Mark
; TITLE OF INVENTION: Modulators of Leukocyte Activation, Hck Compositions and Methods
; FILE REFERENCE: A-71313/RMS/DHR
; CURRENT APPLICATION NUMBER: US/10/193,720
; CURRENT FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1595)
; OTHER INFORMATION:
US-10-193-720-1
```

```

Alignment Scores:
Pred. No.: 2,02e-31 Length: 1924
Score: 360.50 Matches: 77
Percent Similarity: 58.38% Conservative: 31
Best Local Similarity: 41.62% Mismatches: 70
Query Match: 26.64% Indels: 7
DB: 15 Gaps: 2

US-09-939-853A-75 (1-261) x US-10-193-720-1 (1-1924)
Qy 12 ProSerProSerLeuSerSerValGlnGlyProValThrMetGluAlaGlu 31
Db 195 CCGGGCCCTAATAGCCACACAGC-----AACACACACAGGAATCAGGAGCGGC 245
Qy 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeu 51
Db 246 TCTGAGGACATCATCGTGGTTCCTGTATGATTCAGAGCCATTCACACAGAACCTC 305
Qy 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrVal 71
Db 306 AGCTTCAGAGGGGACAGATGGTGTCTTAGAGGAATCCGGGAGTGTGGAGGCT 365
Qy 72 LeuSerGluValSerClyArgGluTyraAsnIleProSerValHisValGlyLysVal--- 90
Db 366 CGATCCCTGGCCACCCGGAAGGAGGCTACATCCCAAGCAACTGTGCGCGGCTGAC 425
Qy 91 -----SerHisGlyTrpLeuTyrgluGlyLeuSerArgGluLysAlaGluLeu 107
Db 426 TCTCTGAGACAGAGGAGTGGTTTCAAGGCGCATCAGCCGGAAGGACGAGAGGCCAA 485
Qy 108 LeuLeuLeuProGlyAsnProGlyAlaPheLeuIleArgGluSerGlnThrArg 127
Db 486 CTGCTGGCTCCCGGCAACATGCTGGGCTCCTTCATGATCCGGGATAGGAGACCACTAA 545
Qy 128 GlySerTyrsSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147
Db 546 GGAAGCTACTCTTGTCCGTGGAGACTACGACCTCGCGGAGAGATACCGTGAACAT 605
Qy 148 TyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrlleSerProArgLeuThrPhePro 167
Db 606 TACAAGATCCGGACCCCTGACACACGGGGCTTCTACATATCCCCCGAAGACCTTCAGC 665
Qy 168 SerLeuGlnAlaLeuValAspHisTyrsSerGluLeuAlaAspAspIleCysCysLeuLeu 187
Db 666 ATCTTCAGAGCTGTGTGACCACTACAGAAGGGGAGACGCGGCTCTGCAAGAAACTG 725
Qy 188 LysGluProCysVal 192
Db 726 TCGGTGCCCTGCATG 740

RESULT 11
US-09-954-456-1983
; Sequence 1983, Application US/09954456
; Patent No. US20030115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26

```

```

; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1983
; LENGTH: 2015
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1983

Alignment Scores:
Pred. No.: 2,16e-31 Length: 2015
Score: 360.50 Matches: 77
Percent Similarity: 58.38% Conservative: 31
Best Local Similarity: 41.62% Mismatches: 70
Query Match: 26.64% Indels: 7
DB: 9 Gaps: 2

US-09-939-853A-75 (1-261) x US-09-954-456-1983 (1-2015)
Qy 12 ProSerProSerLeuSerSerValGlnGlyProValThrMetGluAlaGlu 31
Db 286 CCGGGCCCTAATAGCCACACAGC-----AACACACACAGGAATCAGGAGCGGC 336
Qy 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeu 51
Db 337 TCTGAGGACATCATCGTGGTTCCTGTATGATTCAGAGGCCATTCACCACGAAGACCTC 396
Qy 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrVal 71
Db 397 AGTTCAGAGGGGAGGACAGATGGTGTCTAGAGGAATCCGGGAGTGTGGAGGCT 456
Qy 72 LeuSerGluValSerGlyArgGluTyraAsnIleProSerValHisValGlyLysVal--- 90
Db 457 CGATCCCTGGCCACCCGGAAGGAGGCTACATCCCAAGCAACTGTGCGCGGCTGAC 516
Qy 91 -----SerHisGlyTrpLeuTyrgluGlyLeuSerArgGluLysAlaGluLeu 107
Db 517 TCTCTGAGACAGAGGAGTGGTTTTCAGGGGATCAGCCGGAAGGACGAGAGCCCAA 576
Qy 108 LeuLeuLeuProGlyAsnProGlyAlaPheLeuIleArgGluSerGlnThrArgArg 127
Db 577 CTGCTGGCTCCCGCAACATGCTGGGCTCCTTCATGATCCGGGATAGGAGACCACTAAA 636
Qy 128 GlySerTyrsSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147
Db 637 GGAAGCTACTCTTGTCCGTGGAGACTACGACCTCGCGGAGGAGATACCGTGAACAT 696
Qy 148 TyrArgIleHisCysLeuAspAsnGlyTrpLeuTyrlleSerProArgLeuThrPhePro 167
Db 697 TACAAGATCCGGACCCCTGACACACGGGGCTTCTACATATCCCCCGAAGACCTTCAGC 756
Qy 168 SerLeuGlnAlaLeuValAspHisTyrsSerGluLeuAlaAspAspIleCysCysLeuLeu 187
Db 757 ACTCTGAGGAGCTGTGTGACCACTACAAGAGGGGAGACGCGGCTCTGCCAGAAACTG 816
Qy 188 LysGluProCysVal 192
Db 817 TCGGTGCCCTGCATG 831

RESULT 12
US-10-007-010-3
; Sequence 3, Application US/10007010
; Publication No. US20030125275A1
; GENERAL INFORMATION:
; APPLICANT: Alexander H. Borchers
; APPLICANT: Kenneth W. Dobie

```

; TITLE OF INVENTION: ANTISENSE MODULATION OF HCK EXPRESSION  
 ; FILE REFERENCE: RTS-0345  
 ; CURRENT APPLICATION NUMBER: US/10/007.010  
 ; CURRENT FILING DATE: 2001-12-04  
 ; NUMBER OF SEQ ID NOS: 87  
 ; SEQ ID NO 3  
 ; LENGTH: 2015  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (169)...(1686)  
 US-10-007-010-3

Alignment Scores:  
 Pred. No.: 2,16e-31 Length: 2015  
 Score: 360.50 Matches: 77  
 Percent Similarity: 58.38% Conservative: 31  
 Best Local Similarity: 41.62% Mismatches: 70  
 Query Match: 26.64% Indels: 7  
 DB: 14 Gaps: 2

US-09-939-853A-75 (1-261) x US-10-007-010-3 (1-2015)

QY 12 ProSerProSerLeuSerSerSerValGlnGlyGlnGlyProValThrMetGluAlaGlu 31  
 Db 286 CCGGGGCTAATAGCCCAACAGC-----AACACACACAGGAATCAGGAGGCGAGC 336  
 QY 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeu 51  
 Db 337 TCTGAGACATCATCGTGGTTCCTGTATGATACGAGCAATCACCACGAGACCTC 396  
 QY 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrVal 71  
 Db 397 AGCTTCCAGAGGGGACCATGGTGTCTAGAGGAATCCGGGAGTGTGGAAGGCT 456  
 QY 72 LeuSerGluValSerGlyArgGluTyAsnIleProSerValHisValGlyLysVal--- 90  
 Db 457 CGATCCCTGCCACCCGAGAGGAGGGGTACATCCCAAGCAACTATGTCGCCCGCTTGAC 516  
 QY 91 -----SerHisGlyTrpLeuTyGluGlyLeuSerArgGluLysAlaGluLeu 107  
 Db 517 TCTCTGGAGACAGAGGAGTGTGTTTCAAGGGCATCAGCCGGAAGACGAGAGCCCAA 576  
 QY 108 LeuLeuLeuProGlyAsnProGlyAlaPheLeuIleArgLeuSerGlnThrArgArg 127  
 Db 577 CTGCTGCTCCCGCAACATGCTGGGCTCCTTCATGATCCGGATAGCGAGACCACTAAA 636  
 QY 128 GlySerTySerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147  
 Db 637 GGAAGCTACTTCTTGTCCGCGAGACTACGACCTCGGAGGAGATACCGTGAACAT 696  
 QY 148 TyArgIleHisCysLeuAspAsnGlyTrpLeuTyIleSerProArgLeuThrPhePro 167  
 Db 697 TACAAGATCCGACCCCTGGCAACCGGGGCTTCTACATATCCCCCGAAGACCTTCAG 756  
 QY 168 SerLeuGlnAlaLeuValAspHisTySerGluLeuAlaAspIleCysCysLeuLeu 187  
 Db 757 ACTCTGAGAGCTGGTGGACCACTACAGAGGGGAACGACGGGCTCTGCCAGAAACTG 816  
 QY 188 LysGluProCysVal 192  
 Db 817 TCGGTGCCCTGCATG 831

RESULT 13  
 US-10-252-157-140  
 ; Sequence 140, Application US/10252157  
 ; Publication No. US20030190640A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Faris, Mary  
 ; APPLICANT: Pearson, Cecelia I.  
 ; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER  
 ; FILE REFERENCE: PA-0027-1 US

; CURRENT APPLICATION NUMBER: US/10/252.157  
 ; CURRENT FILING DATE: 2002-10-01  
 ; PRIOR APPLICATION NUMBER: 60/295,048  
 ; PRIOR FILING DATE: 2001-05-31  
 ; NUMBER OF SEQ ID NOS: 501  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 140  
 ; LENGTH: 2341  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20030190640A1 411296.2  
 US-10-252-157-140

Alignment Scores:  
 Pred. No.: 2,65e-31 Length: 2341  
 Score: 360.50 Matches: 77  
 Percent Similarity: 58.38% Conservative: 31  
 Best Local Similarity: 41.62% Mismatches: 70  
 Query Match: 26.64% Indels: 7  
 DB: 14 Gaps: 2

US-09-939-853A-75 (1-261) x US-10-252-157-140 (1-2341)

QY 12 ProSerProSerLeuSerSerSerValGlnGlyGlnGlyProValThrMetGluAlaGlu 31  
 Db 592 CCGGGGCTAATAGCCCAACAGC-----AACACACACAGGAATCAGGAGGCGAGC 642  
 QY 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeu 51  
 Db 643 TCTGAGACATCATCGTGGTTCCTGTATGATACGAGCAATCACCACGAGACCTC 702  
 QY 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrVal 71  
 Db 703 AGCTTCCAGAGGGGACCATGGTGTCTAGAGGAATCCGGGAGTGTGGAAGGCT 762  
 QY 72 LeuSerGluValSerGlyArgGluTyAsnIleProSerValHisValGlyLysVal--- 90  
 Db 763 CGATCCCTGCCACCCGAGAGGAGGGGTACATCCCAAGCAACTATGTCGCCCGCTTGAC 822  
 QY 91 -----SerHisGlyTrpLeuTyGluGlyLeuSerArgGluLysAlaGluLeu 107  
 Db 823 TCTCTGGAGACAGAGGAGTGTGTTTCAAGGGCATCAGCCGGAAGACGAGAGCCCAA 882  
 QY 108 LeuLeuLeuProGlyAsnProGlyAlaPheLeuIleArgLeuSerGlnThrArgArg 127  
 Db 883 CTGCTGCTCCCGCAACATGCTGGGCTCCTTCATGATCCGGATAGCGAGACCACTAAA 942  
 QY 128 GlySerTySerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147  
 Db 943 GGAAGCTACTTCTTGTCCGCGAGACTACGACCTCGGAGGAGATACCGTGAACAT 1002  
 QY 148 TyArgIleHisCysLeuAspAsnGlyTrpLeuTyIleSerProArgLeuThrPhePro 167  
 Db 1003 TACAAGATCCGACCCCTGGCAACCGGGGCTTCTACATATCCCCCGAAGACCTTCAGC 1062  
 QY 168 SerLeuGlnAlaLeuValAspHisTySerGluLeuAlaAspIleCysCysLeuLeu 187  
 Db 1063 ACTCTGAGAGCTGGTGGACCACTACAGAGGGGAACGACGGGCTCTGCCAGAAACTG 1122  
 QY 188 LysGluProCysVal 192  
 Db 1123 TCGGTGCCCTGCATG 1137

RESULT 14  
 US-10-062-674-2038  
 ; Sequence 2038, Application US/10062674  
 ; Publication No. US2004000559A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.  
 ; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
 ; FILE REFERENCE: PA-0026-1 CIP

; CURRENT APPLICATION NUMBER: US/10/062,674  
 ; CURRENT FILING DATE: 2002-01-30  
 ; PRIOR APPLICATION NUMBER: US 09/625,102  
 ; PRIOR FILING DATE: 2000-07-24  
 ; NUMBER OF SEQ ID NOS: 2217  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 2038  
 ; LENGTH: 2343  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20040005559A1 411296.5  
 US-10-062-674-2038

Alignment Scores:  
 Pred. No.: 2,66e-31 Length: 2343  
 Score: 360.50 Matches: 77  
 Percent Similarity: 58.38% Conservative: 31  
 Best Local Similarity: 41.62% Mismatches: 70  
 Query Match: 26.64% Indels: 7  
 DB: 15 Gaps: 2

US-09-939-853A-75 (1-261) x US-10-062-674-2038 (1-2343)

QY 12 ProSerProSerLeuSerSerValGlnGlnGlyProValThrMetGluAlaGlu 31  
 Db 592 CCGGGGCTAATAGCCACACAGC-----AACACACACGGAATCAGGGAGCGAGC 642  
 QY 32 ArgSerLysAlaThrAlaValAlaLeuGlySerPheProAlaGlyProAlaGluLeu 51  
 Db 643 TCTGAGACATCATCGTGGTTCCTGTATGATACGAGCCATTCACACGAGACCTC 702  
 QY 52 SerLeuArgLeuGlyGluProLeuThrIleValSerGluAspGlyAspTrpThrVal 71  
 Db 703 AGTTCAGAGGGGACACAGATGCTGCTAGAGGATCCGGGAGTGTGGAGGCT 762  
 QY 72 LeuSerGluValSerGlyArgGlyTyrAsnIleProSerValHisValGlyLysVal--- 90  
 Db 763 CGATCCCTGCCCGGAGAGGGGTACATCCCAAGCAACTATGTCGCCGCGGTGAC 822  
 QY 91 -----SerHisGlyTyrLeuTyrGlyLeuSerArgGlyLysAlaGluLeu 107  
 Db 823 TCTCTGAGACAGAGGATGTTTTCAGGGGATCAGCCGAGAGACGACGACCCAA 882  
 QY 108 LeuLeuLeuProGlyAenProGlyAlaPheLeuLeuArgGluSerGlnThrArgG 127  
 Db 883 CTGCTGCTCCCGCAACATGCTGGGCTCTTCATGATCCGGATAGCGAGACCACTAA 942  
 QY 128 GlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTrpAspArgIleArgHis 147  
 Db 943 GGAAGTACTCTCTGCTCGGCGAGACTACGACCTCGCAGGAGATACCGTGAACAT 1002  
 QY 148 TyrArgIleHisCysLeuAspAenGlyTyrLeuTyrIleSerProArgLeuThrPhePro 167  
 Db 1003 TACAAGATCCGACCTGGACAACGGGGCTTCTACATATCCCGGAGACCTTCAGC 1062  
 QY 168 SerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeu 187  
 Db 1063 ACTCTCAGAGAGTGTGGACCACTACAAGAGGGGAGCAGCGGGCTCTGCCAGAACTG 1122  
 QY 188 LysGluProCysVal 192  
 Db 1123 TCGGTGCCCTGCATG 1137

RESULT 15

; US-09-917-800A-1611  
 ; Sequence 1611, Application US/09917800A  
 ; Patent No. US20020119462A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mendrick, Donna  
 ; APPLICANT: Porter, Mark  
 ; APPLICANT: Johnson, Kory

; APPLICANT: Castle, Arthur  
 ; APPLICANT: Blashoff, Michael  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Molecular Toxicology Modeling  
 ; FILE REFERENCE: 4921-5038-US  
 ; CURRENT APPLICATION NUMBER: US/09/917,800A  
 ; CURRENT FILING DATE: 2001-07-31  
 ; PRIOR APPLICATION NUMBER: US 60/222,040  
 ; PRIOR FILING DATE: 2000-07-31  
 ; PRIOR APPLICATION NUMBER: US 60/222,880  
 ; PRIOR FILING DATE: 2000-11-02  
 ; PRIOR APPLICATION NUMBER: US 60/290,029  
 ; PRIOR FILING DATE: 2001-05-11  
 ; PRIOR APPLICATION NUMBER: US 60/290,645  
 ; PRIOR FILING DATE: 2001-05-15  
 ; PRIOR APPLICATION NUMBER: US 60/292,336  
 ; PRIOR FILING DATE: 2001-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/295,798  
 ; PRIOR FILING DATE: 2001-06-06  
 ; PRIOR APPLICATION NUMBER: US 60/297,457  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,884  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/303,459  
 ; PRIOR FILING DATE: 2001-07-09  
 ; NUMBER OF SEQ ID NOS: 1740  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1611  
 ; LENGTH: 1911  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_013185  
 US-09-917-800A-1611

Alignment Scores:  
 Pred. No.: 3,41e-31 Length: 1911  
 Score: 358.50 Matches: 93  
 Percent Similarity: 52.22% Conservative: 48  
 Best Local Similarity: 34.44% Mismatches: 94  
 Query Match: 26.50% Indels: 35  
 DB: 9 Gaps: 6

US-09-939-853A-75 (1-261) x US-09-917-800A-1611 (1-1911)

QY 1 MetGlySerLeuProSerArg-----ArgLysSerLeuProSerProSerLeuSerSer 18  
 Db 184 ATGGGATGTGTGAAGTCCAGGTCTCCGAGAGGAGCAAGGCTCAAAAATAGAGCCA 243  
 QY 19 SerValGlnGlyGlnGlyProValThrMetGlu----- 29  
 Db 244 AATGCCAACCAAGAAAGGCGCTGTGTATGTCCGGATCCACGTCCTCCCTAAGAGCTGGGA 303  
 QY 30 -----AlaGluArgSerLysAlaThr---Ala 37  
 Db 304 CCGAACAGCATCAACACCTGCCCCCGGGTTCGTGGAGGGCTCTGAGACACCATTTGTG 363  
 QY 38 ValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGlu 57  
 Db 364 GTCGCACTGTACGACTATGAGGCCATTCCACGCTGAAGACCTCAGCTTCCAGAGGGAGAC 423  
 QY 58 ProLeuThrIleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGly 77  
 Db 424 CAGATGTGTCTCTGGAGGAGTCTGGGGAGTGTGGAGCGCCGCTTCCCTGGCTACCAAG 483  
 QY 78 ArgGluTyrAsnIleProSerValHisValGlyLysVal-----SerHisGly 93  
 Db 484 AAGAAGGCTATATCCCAAGCAATATGTAGTCGAGTAACTCTTTGGAGACTGAGGAG 543  
 QY 94 TrpLeuTyrGluGlyLeuSerArgGluLysAlaGluLeuLeuLeuProGlyAsn 113  
 Db 544 TGGTCTTCAAGGGTATCAGCCGGAAGGATGACAGCGCCACCTGTGCTGCCCGGAGAC 603

```

QY 114 ProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTySerLeuSer 133
Db 604 ATGCTGGGCTCCTTCATGATCCGGACAGTGAGACCAACCAAGGGAGCTACTCACTTTCT 663
QY 134 ValArgLeuSerArgProAlaSerTrpAspArgIleArgHisTyArgIleHisCysLeu 153
Db 664 GTTCGAGACTTTGACCCGCCAGCAGGAGACACGGTGAGCATTATAAATCCGGACACTG 723
QY 154 AspAsnGlyTrpLeuTyIleSerProArgLeuThrPheProSerLeuGlnAlaLeuVal 173
Db 724 GACAGTGGAGGGTTCTACATCTCCGAGGAGCACCTTCAGCAGCCTGCGAGAACTTGTGTC 783
QY 174 AspHisTySerGluLeuAlaAspIleCysCysLeuLeuLysGluProCysVal-Ie 193
Db 784 GTCCACTACAAGAGGGAGAGATGGGCTCTGCCAGAGCTGTCAGTGCCTGTGTGTCT 843
QY 193 uGln-----ArgAlaGlyProLeuProGlyLysAspIleProLeuProValTh 209
Db 844 CCGAAACCCAGAACCCATGGGAGAAAGATGCTGGGAGATTCCTCGAGAATCCCTGCAG 903
QY 209 rValGlnArgThr-----ProLeuAsn-TTpLysGluLeuAspSerSerLeuLeuPheS 227
Db 904 ATGGAGAGAAACTGGGAGCCGGCAGTTTGGAGAGTGTGGATGGCCACCTACAACAAG 963
QY 227 erGluAlaAlaThrGlyGluGluSer 235
Db 964 CACACCAAAGTGCGGTGAAGACAAT 989

```

Search completed: February 20, 2004, 12:38:07  
Job time : 370 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:18 ; Search time 87.562 Seconds  
(without alignments)  
9899.970 Million cell updates/sec

Title: US-09-939-853A-140

Perfect score: 20

Sequence: 1 ctggacaggttagggcttgg 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*\*

1: gb\_ba:\*\*

2: gb\_htg:\*\*

3: gb\_in:\*\*

4: gb\_om:\*\*

5: gb\_ov:\*\*

6: gb\_pat:\*\*

7: gb\_ph:\*\*

8: gb\_pi:\*\*

9: gb\_pr:\*\*

10: gb\_ro:\*\*

11: gb\_sts:\*\*

12: gb\_sy:\*\*

13: gb\_un:\*\*

14: gb\_vi:\*\*

15: em\_ba:\*\*

16: em\_fun:\*\*

17: em\_hum:\*\*

18: em\_in:\*\*

19: em\_mu:\*\*

20: em\_om:\*\*

21: em\_or:\*\*

22: em\_ov:\*\*

23: em\_pat:\*\*

24: em\_ph:\*\*

25: em\_pi:\*\*

26: em\_ro:\*\*

27: em\_sts:\*\*

28: em\_un:\*\*

29: em\_vi:\*\*

30: em\_htg\_hum:\*\*

31: em\_htg\_inv:\*\*

32: em\_htg\_other:\*\*

33: em\_htg\_mus:\*\*

34: em\_htg\_pln:\*\*

35: em\_htg\_rnd:\*\*

36: em\_htg\_vrt:\*\*

37: em\_htg\_vrt:\*\*

38: em\_sy:\*\*

39: em\_htgo\_hum:\*\*

40: em\_htgo\_mus:\*\*

41: em\_htgo\_other:\*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	20	100.0	20	6	AX443199	AX443199 Sequence
2	20	100.0	1183	6	AX443133	AX443133 Sequence
3	20	100.0	1183	6	AX443135	AX443135 Sequence
4	20	100.0	2538	9	BC042041	BC042041 Homo sapi
5	20	100.0	2567	6	AX452880	AX452880 Sequence
6	20	100.0	2788	6	AX780857	AX780857 Sequence
7	20	100.0	6741	9	HS460038	AL031862 Human DNA
8	20	100.0	145833	2	AC026539	AC026539 Homo sapi
9	17.4	87.0	80828	9	AC093247	AC093247 Homo sapi
10	17.4	87.0	100300	2	AC114970	AC114970 Homo sapi
11	17.4	87.0	108742	2	AC023817	AC023817 Homo sapi
12	17.4	87.0	110000	2	AC139485_1	Continuation (2 of
13	17.4	87.0	115702	9	AC132801	AC132801 Homo sapi
14	17.4	87.0	129218	2	AC138851	AC138851 Homo sapi
15	17.4	87.0	133181	2	AC138930	AC138930 Homo sapi
16	17.4	87.0	140410	2	AC139483	AC139483 Homo sapi
17	17.4	87.0	145493	2	AC138839	AC138839 Homo sapi
18	17.4	87.0	146061	9	AC118459	AC118459 Homo sapi
19	17.4	87.0	146175	2	AL161433	AL161433 Homo sapi
20	17.4	87.0	147840	2	AC138939	AC138939 Homo sapi
21	17.4	87.0	149397	9	AC138827	AC138827 Homo sapi
22	17.4	87.0	152445	9	AC138410	AC138410 Homo sapi
23	17.4	87.0	152782	9	AL133255	AL133255 Human DNA
24	17.4	87.0	154803	9	AC131392	AC131392 Homo sapi
25	17.4	87.0	155537	2	AC139282	AC139282 Homo sapi
26	17.4	87.0	156527	2	AC138968	AC138968 Homo sapi
27	17.4	87.0	158755	2	AC145032	AC145032 Homo sapi
28	17.4	87.0	159184	2	AC139484	AC139484 Homo sapi
29	17.4	87.0	160994	2	AC044800	AC044800 Homo sapi
30	17.4	87.0	161000	2	AC145100	AC145100 Homo sapi
31	17.4	87.0	161279	2	AC139508	AC139508 Homo sapi
32	17.4	87.0	161460	9	AC093243	AC093243 Homo sapi
33	17.4	87.0	161460	9	AL157879	AL157879 Human DNA
34	17.4	87.0	161671	9	AC140175	AC140175 Homo sapi
35	17.4	87.0	161893	9	AC137781	AC137781 Homo sapi
36	17.4	87.0	161957	2	AC138909	AC138909 Homo sapi
37	17.4	87.0	162703	9	AC132802	AC132802 Homo sapi
38	17.4	87.0	165211	2	AC145128	AC145128 Homo sapi
39	17.4	87.0	166615	9	AC138928	AC138928 Homo sapi
40	17.4	87.0	167663	2	AC145135	AC145135 Homo sapi
41	17.4	87.0	168547	2	AC138947	AC138947 Homo sapi
42	17.4	87.0	168814	9	AC010237	AC010237 Homo sapi
43	17.4	87.0	169082	2	AC138933	AC138933 Homo sapi
44	17.4	87.0	170336	2	AC011244	AC011244 Homo sapi
45	17.4	87.0	171363	9	AC145129	AC145129 Homo sapi

# ALIGNMENTS

RESULT 1  
AX443199  
LOCUS AX443199  
DEFINITION Sequence 140 from Patent WO0216599.  
ACCESSION AX443199  
VERSION AX443199.1 GI:21690594  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1  
AUTHORS Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R.,  
Shinkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E.,  
Topper, J.N. and Yang, R.B.  
TITLE Proteins and nucleic acids encoding same

AX443199 20 bp DNA linear PAT 02-JUL-2002

JOURNAL Patent: WO 0216599-A 140 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES Location/Qualifiers

source  
1..20  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="oligonucleotide primer"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20  
|||||  
Db 1 CTGGACAGGTTAGGCGTTTG 20

RESULT 2  
LOCUS AX443133 1183 bp DNA linear PAT 02-JUL-2002  
DEFINITION Sequence 74 from Patent WO0216599.  
ACCESSION AX443133  
KEYWORDS AX443133.1 GI:21690555  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,  
Shinkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,  
Topper,J.N. and Yang,R.B.  
TITLE Proteins and nucleic acids encoding same

JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES source  
1..1183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20  
|||||  
Db 301 CTGGACAGGTTAGGCGTTTG 282

RESULT 3  
LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002  
DEFINITION Sequence 76 from Patent WO0216599.  
ACCESSION AX443135  
VERSION AX443135.1 GI:21690556  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,  
Shinkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,  
Topper,J.N. and Yang,R.B.  
TITLE Proteins and nucleic acids encoding same

JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;  
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)  
FEATURES source  
1..1183

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20  
|||||  
Db 883 CTGGACAGGTTAGGCGTTTG 902

## RESULT 4

BC042041/c  
LOCUS BC042041 2538 bp mRNA linear PRI 07-OCT-2003  
DEFINITION Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (CDNA  
clone MGC:49845 IMAGE:4429896), complete cds.

ACCESSION BC042041  
VERSION BC042041.1 GI:27469842  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2538)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Mialaby,S.J., Bosak,S.A., McEwan,P.J.,  
Mokernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S.A., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,  
Schurch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

## TITLE

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mdc@paxil.stanford.edu](mailto:mdc@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 88 Row: a Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28416422.

```

FEATURES
  source
    1. .2538
      /location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="MGC:49845 IMAGE:4429896"
        /tissue_type="Prostate, adenocarcinoma."
        /clone_lib="NIH_MGC_91"
        /lab_host="DH10B"
        /note="Vector: pCMV-SPORT6"
      1. .2538
        /gene="SLA2"
        /note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845,
        SLAP2"
        /db_xref="LocusID:84174"
        /db_xref="MIM:60577"
        363. .1148
          /codon_start=1
          /product="Src-like-adaptor 2, isoform a"
          /protein_id="AAH42041.1"
          /db_xref="GI:27459843"
          /db_xref="LocusID:84174"
          /translation="MGSLLPERRKSLPSPSSVQGGPVVMEARSKATAVALGSPF
          AGPAELSLRLGELTVSDGDWTVLSEVGRYNPVSHVAKVSHGWLNGLSRE
          KAEELLPGNPGGAFIRSGYSLSVLRSPASWDRIHRVIRHGLDNGWLYI
          SPLTTPSLOALVDHYSELADICCLKPCVLQAGPLPGKIDPLPVTVQPTPLNKK
          ELSSLLPSEARTEGESLLSEGLRESLSFVSLNDEAVSLDDA"
        468. .632
          /note="SH3; Region: SH3 domain. SH3 (Src homology 3)
          domains are often indicative of a protein involved in
          signal transduction related to cytoskeletal organization.
          First described in the Src cytoplasmic tyrosine kinase.
          The structure is a partly opened beta barrel"
          /db_xref="CDD:pfam00018"
        542. .890
          /note="SH2; Region: SH2 domain"
          /db_xref="CDD:pfam00017"

  CDS
    1. .2538
      /product="Src-like-adaptor 2, isoform a"
      /protein_id="AAH42041.1"
      /db_xref="GI:27459843"
      /db_xref="LocusID:84174"
      /translation="MGSLLPERRKSLPSPSSVQGGPVVMEARSKATAVALGSPF
      AGPAELSLRLGELTVSDGDWTVLSEVGRYNPVSHVAKVSHGWLNGLSRE
      KAEELLPGNPGGAFIRSGYSLSVLRSPASWDRIHRVIRHGLDNGWLYI
      SPLTTPSLOALVDHYSELADICCLKPCVLQAGPLPGKIDPLPVTVQPTPLNKK
      ELSSLLPSEARTEGESLLSEGLRESLSFVSLNDEAVSLDDA"
      468. .632
        /note="SH3; Region: SH3 domain. SH3 (Src homology 3)
        domains are often indicative of a protein involved in
        signal transduction related to cytoskeletal organization.
        First described in the Src cytoplasmic tyrosine kinase.
        The structure is a partly opened beta barrel"
        /db_xref="CDD:pfam00018"
      542. .890
        /note="SH2; Region: SH2 domain"
        /db_xref="CDD:pfam00017"

  misc_feature
    542. .890
      /note="SH2; Region: SH2 domain"
      /db_xref="CDD:pfam00017"

  misc_feature
    542. .890
      /note="SH2; Region: SH2 domain"
      /db_xref="CDD:pfam00017"

  ORIGIN
    Query Match 100.0%; Score 20; DB 9; Length 2538;
    Best Local Similarity 100.0%; Pred. No. 4.4;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    Qy 1 CTGGACAGGTTAGGCGTTTG 20
    Db 266 CTGGACAGGTTAGGCGTTTG 247

  RESULT 5
  AX452880/c
  LOCUS 2567 bp DNA linear PAT 06-JUL-2002
  DEFINITION Sequence 1 from Patent WO0242457.
  ACCESSION AX452880
  VERSION AX452880.1 GI:21712520
  KEYWORDS Homo sapiens (human)
  SOURCE Homo sapiens
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villar, J.J. and
    Kanner, S.B.
    Cloning and expression of human slap-2: a novel sh2/sh3
    domain-containing human slap homologue having immune cell-specific
    expression
    Patent: WO 0242457-A 1 30-MAY-2002;
    Bristol-Myers Squibb Co. (US)
  JOURNAL
    Location/Qualifiers
  FEATURES
    source
      1. .2567
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"

```

/db\_xref="taxon:9606"

```

  ORIGIN
    Query Match 100.0%; Score 20; DB 6; Length 2567;
    Best Local Similarity 100.0%; Pred. No. 4.4;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    Qy 1 CTGGACAGGTTAGGCGTTTG 20
    Db 318 CTGGACAGGTTAGGCGTTTG 299

  RESULT 6
  AX780857/c
  LOCUS 2788 bp DNA linear PAT 14-JUL-2003
  DEFINITION Sequence 3014 from Patent WO03039443.
  ACCESSION AX780857
  VERSION AX780857.1 GI:32697851
  KEYWORDS Homo sapiens (human)
  SOURCE Homo sapiens
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 Haeflrich, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
    Dugas, M., Ellis, R., Brots, B. and Mergenthaler, S.
    Novel genetic markers for leukemias
    Patent: WO 03039443-A 3014 15-MAY-2003;
    Deutsches Krebsforschungszentrum (DE);
    Ludwig-Maximilian-Universitaet Muenchen (DE); Haeflrich, Torsten,
    PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
  FEATURES
    source
      1. .2788
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

  ORIGIN
    Query Match 100.0%; Score 20; DB 6; Length 2788;
    Best Local Similarity 100.0%; Pred. No. 4.4;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    Qy 1 CTGGACAGGTTAGGCGTTTG 20
    Db 290 CTGGACAGGTTAGGCGTTTG 271

  RESULT 7
  HS460J8
  LOCUS 66741 bp DNA linear PRI 23-JUL-2001
  DEFINITION Human DNA sequence from clone RP3-460J8 on chromosome
  20q11.21-11.23 Contains the 3' end of the gene for a novel protein
  similar to N-myc downstream regulated (NDRG1) the 5' end of a gene
  encoding a novel protein tyrosine kinase, ESTs, STSS and GSSs,
  complete sequence.
  AL031662
  AL031662.26 GI:9716901
  VERSION HTG; NDRG1; SH2 domain.
  KEYWORDS Homo sapiens (human)
  SOURCE Homo sapiens
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 66741)
    Skuce, C.
  AUTHORS Direct Submission
  TITLE Submitted (23-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
  JOURNAL requests: clonerequest@sanger.ac.uk
  COMMENT
    On Aug 7, 2000 this sequence version replaced gi:6425549.
    During sequence assembly data is compared from overlapping clones.
    Where differences are found these are annotated as variations
    together with a note of the overlapping clone name. Note that the
    variation annotation may not be found in the sequence submission

```

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human

chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr20>

IMPORTANT: This sequence is not the entire insert of clone RP3-460J8 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP3-460J8 is at 6741 in this sequence. The true left end of clone RP5-977B1 is at 41767 in this sequence.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP3-460J8 is from the library RPCI-3 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCVPAC2.

#### FEATURES

##### source

##### Location/Qualifiers

1. .66741

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="20"

/map="G11-21-11.23"

/clone="RP3-460J8"

/clone\_lib="RPCI-3"

/complement(50. .544)

/note="match: GSS: Em:B45150"

/complement(240. .7128)

/gene="dJ977B1.1"

/complement(join(<240. .339,6995. .7128))

/gene="dJ977B1.1"

/product="dJ460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"

/note="match: cDNAs: Em:AK025645"

/match: ESTs: Em:BG178487"

/evidence=not experimental

/complement(join(<240. .339,6995. .7085))

/gene="dJ977B1.1"

/note="Continues in Em:AL050318 as dJ977B1.2"

/codon\_start=1

/evidence=not experimental

/product="dJ460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"

/protein\_id="CAC44645.1"

/db\_xref="GI:15020830"

/db\_xref="GOA:Q9H6Q3"

/db\_xref="SWISS-PROT:Q9H6Q3"

/translation="MGSLPSRRKSLPSLSSVQGGQVPTMEATERSKATAVALGSPFP AGGPALSLRLGLEPTIVS"

536. .565

/note="15 copies 2 mer ag 93% conserved"

2375. .2402

/note="14 copies 2 mer ta 100% conserved"

2783. .3155

/note="match: GSS: Em:AQ807191"

/complement(6983. .7482)

/note="match: GSS: Em:AQ556467"

/complement(5985. .7492)

/note="match: GSS: Em:AQ556478"

7486. .7925

/note="match: STS: Em:HS427J1S"

##### repeat\_region

10615. .10640

/note="13 copies 2 mer tt 92% conserved"

13593. .13961

/note="MER21B repeat: matches 422. .785 of consensus"

14248. .14297

/note="MER21B repeat: matches 374. .422 of consensus"

14594. .14614

/note="MER21B repeat: matches 355. .374 of consensus"

14746. .15078

/note="MER21B repeat: matches 1. .355 of consensus"

/complement(17516)

17519. .17686

/note="match: GSS: Em:AZ067993"

/complement(17521)

complement(17522)

/gene="dJ469A13.3"

complement(join(<17523. .19451,20551. .20589,22110. .22161,22247. .22282,26087. .26134,30796. .30847,30986. .31051,32039. .32142,37091. .37147,46566. .46652,48268. .48328,50163. .50225,53242. .53362,54429. .54534))

/gene="dJ469A13.3"

/product="dJ460J8.1 (continued from dJ469A13.3 in Em:AL132768)"

/note="match: cDNAs: Em:AB033922 Em:U52073 Em:D87953 Em:X92845 Em:AF004162 Em:U60593 Em:AF045564 Em:AB033921 Em:ARL59092 Em:AF147402 Em:M59814

match: ESTs: Em:AA718726 Em:AA039000 Em:AW003952 Em:W89263 Em:AA325826 Em:AI230982 Em:AA162360 Em:AA445016 Em:AV002395 Em:T85147 Em:AI786615 Em:AA113437 Em:AI004026 Em:T88705 Em:AI786673 Em:AI181197 Em:AA764653 Em:AA316771 Em:AU035165 Em:AV002368"

/evidence=not experimental

/complement(17542. .17547)

/gene="dJ469A13.3"

17555. .17811

/note="match: STS: Em:G19945"

/join(19202. .19451,22110. .22282,26087. .26127,30796. .30847)

/note="match: STS: Em:G23762"

complement(join(19270. .19451,20551. .20589,22110. .22161,22247. .22282,26087. .26134,30796. .30847,30986. .31051,32039. .32142,37091. .37147,46566. .46652,48268. .48328,50163. .50225,53242. .53362,54429. .54534))

/gene="dJ469A13.3"

/note="novel protein (FLJ13556) similar to N-myc downstream regulated (NDRG1)

match: proteins: Sw:Q62433 Sw:Q92597 Tr:Q92219 Sw:P97862"

/codon\_start=1

/evidence=not experimental

/product="dJ460J8.1 (continued from dJ469A13.3 in Em:AL132768)"

/protein\_id="CAB65625.1"

/db\_xref="GI:6687781"

/db\_xref="GOA:Q9UGV2"

/db\_xref="SWISS-PROT:Q9UGV2"

/translation="EHD1ETGVVHTVIRGLPKGNRPVILTYHDIHLNHSKCNFAFF NFEDQETIQHFVAVCHVDAPGQCEGAPFPPTGYQVETMDELAEMLPVILTHLSKSLIIGVGAGATILRFPALNFELVGLVNLINVDPCAKGIDNWAASKLSGLITNVVDILLAHFHFGEEQANLDLIQTYRMHIAQDINQIQLFNLSNGRDLIERPILGDNOKSKTKLKCSTLLVGDGNSFAVAVVEVCNSRLNPIITLLKWDCCGLPQVPGKLTAFK YFLQMGVYPPYQLSHLSTESVPSASMTLARSRTHTSSSLGSGESFPFSRSTVSNQSDGTQSCESPVDLDRHQTMEVSC"

complement(24246. .24780)

/gene="dJ469A13.3"

/note="match: GSS: Em:AQ592789"

complement(27980. .28556)

/gene="dJ469A13.3"

/note="match: GSS: Em:AQ308867"

34313. .34420

/note="HY1 repeat: matches 1. .109 of consensus"

34505. .34560

/note="28 copies 2 mer ta 78% conserved"

34564. .34611

##### misc\_feature

complement(27980. .28556)

/note="match: GSS: Em:AQ592789"

complement(27980. .28556)

/gene="dJ469A13.3"

/note="match: GSS: Em:AQ308867"

34313. .34420

/note="HY1 repeat: matches 1. .109 of consensus"

34505. .34560

/note="28 copies 2 mer ta 78% conserved"

34564. .34611

/note="match: STS: Em:HS427J1S"

```

repeat_region /note="24 copies 2 mer ta 79% conserved"
42553..43658
/note="MEX52C repeat: matches 1. 1278 of consensus"
47989..49470
/note="match: GSS: Em:AQ672417"
49391..49480
/note="45 copies 2 mer ta 76% conserved"
52922..53120
/note="match: STS: Em:G04621"
/note="match: STS: Em:G04621"
complement(53426..54032)
/note="match: GSS: Em:AQ314824
match: STS: Em:G55853"
54054..54477
/note="match: GSS: Em:AQ310681"
57509..57838
/note="match: STS: Em:G07504"
complement(57696..58013)
/note="match: STS: Em:G07632 Em:G07634"
58856..59001
/note="73 copies 2 mer at 83% conserved"
complement(58930..59281)
/note="match: GSS: Em:AQ067563"
59062..59236
/note="TIGER2 repeat: matches 2541. 2714 of consensus"
59738..60495
/note="TIGER2 repeat: matches 1780. 2541 of consensus"
60719..61831
/note="MER11C repeat: matches 1. 1071 of consensus"
62847..63174

```

```

Query Match 100.0%; Score 20; DB 9; Length 66741;
Best Local Similarity 100.0%; Pred. NO. 4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGTAGGGCTTTG 20
Db 11840 CTGCACAGGTAGGGCTTTG 11659

RESULT 8
AC026539/c 145833 bp DNA linear HTG 27-APR-2000
LOCUS Homo sapiens chromosome 20 clone RP11-712N14 map 20, WORKING DRAFT
DEFINITION SEQUENCE, 37 unordered pieces.
ACCESSION AC026539
VERSION AC026539.2 GI:7656813
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145833)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,I., Boukhgalter,B., Brown,A., Burkett,G.,
Campione,A., Castelle,A., Choepe,Y., Colangelo,M., Collins,S.,
Collipetro,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lakocque,K., Lamazares,R., Landers,T., Lenocsky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGuck,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,

```

Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 27, 2000 this sequence version replaced gi:7283243.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center -----

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIER  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information -----  
 Center project name: L7115  
 Center clone name: 712.N.14  
 ----- Summary Statistics -----

Sequencing vector: M13; M7815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 125577 bases at least Q40  
 Consensus quality: 135703 bases at least Q30  
 Consensus quality: 139593 bases at least Q20  
 Insert size: 182000; agarose-fp  
 Quality size: 142233; sum-of-contigs  
 Quality coverage: 2.6 in Q20 bases; agarose-fp  
 Quality coverage: 3.3 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 37 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 1214: contig of 1214 bp in length  
 \* 1215 1314: gap of 100 bp  
 \* 1315 2673: contig of 1359 bp in length  
 \* 2674 2773: gap of 100 bp  
 \* 2774 4520: contig of 1747 bp in length  
 \* 4521 4620: gap of 100 bp  
 \* 4621 5961: contig of 1341 bp in length  
 \* 5962 6061: gap of 100 bp  
 \* 6062 7719: contig of 1658 bp in length  
 \* 7720 7819: gap of 100 bp  
 \* 7820 9799: contig of 1980 bp in length  
 \* 9800 9899: gap of 100 bp  
 \* 9900 11434: contig of 1535 bp in length  
 \* 11435 11534: gap of 100 bp  
 \* 11535 14382: contig of 2848 bp in length  
 \* 14383 14482: gap of 100 bp  
 \* 14483 17148: contig of 2666 bp in length  
 \* 17149 17248: gap of 100 bp  
 \* 17249 19239: contig of 1991 bp in length  
 \* 19240 19339: gap of 100 bp  
 \* 19340 21102: contig of 1763 bp in length  
 \* 21103 21202: gap of 100 bp  
 \* 21203 23371: contig of 2169 bp in length  
 \* 23372 23471: gap of 100 bp  
 \* 23472 25782: contig of 2311 bp in length  
 \* 25783 25882: gap of 100 bp  
 \* 25883 28824: contig of 2942 bp in length  
 \* 28825 28924: gap of 100 bp  
 \* 28925 31619: contig of 2695 bp in length  
 \* 31620 31719: gap of 100 bp  
 \* 31720 34679: contig of 2960 bp in length  
 \* 34680 34779: gap of 100 bp  
 \* 34780 37527: contig of 2748 bp in length

```
* 37528 37627: gap of 100 bp
* 37628 40146: contig of 2519 bp in length
* 40147 40246: gap of 100 bp
* 40247 43743: contig of 3497 bp in length
* 43744 43843: gap of 100 bp
* 43844 47133: contig of 3290 bp in length
* 47134 47233: gap of 100 bp
* 47234 51023: contig of 3790 bp in length
* 51024 51123: gap of 100 bp
* 51124 54935: contig of 3812 bp in length
* 54936 55035: gap of 100 bp
* 55036 59553: contig of 4518 bp in length
* 59554 59653: gap of 100 bp
* 59654 64428: contig of 4775 bp in length
* 64429 64528: gap of 100 bp
* 64529 69211: contig of 4583 bp in length
* 69212 69311: gap of 100 bp
* 69312 72901: contig of 3590 bp in length
* 72902 73001: gap of 100 bp
* 73002 76724: contig of 3723 bp in length
* 76725 76824: gap of 100 bp
* 76825 81179: contig of 4355 bp in length
* 81180 81279: gap of 100 bp
* 81280 87009: contig of 5730 bp in length
* 87010 87109: gap of 100 bp
* 87110 90855: contig of 3746 bp in length
* 90856 90955: gap of 100 bp
* 90956 96520: contig of 5565 bp in length
* 96521 96620: gap of 100 bp
* 96621 102321: contig of 5701 bp in length
* 102322 102421: gap of 100 bp
* 102422 108293: contig of 5872 bp in length
* 108294 108393: gap of 100 bp
* 108394 116689: contig of 8296 bp in length
* 116690 116789: gap of 100 bp
* 116790 125264: contig of 8475 bp in length
* 125265 125364: gap of 100 bp
* 125365 136354: contig of 10990 bp in length
* 136355 136454: gap of 100 bp
* 136455 145833: contig of 9379 bp in length.

FEATURES
    source
        Location/Qualifiers
            1..145833
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /map="20"
                /chromosome="20"
                /clone="RP11-712N14"
                /clone.lib="RPC1-11 Human Male BAC"
            1..1214
                /note="assembly_fragment"
            1315..2673
                /note="assembly_fragment"
            2774..4520
                /note="assembly_fragment"
            4621..5961
                /note="assembly_fragment"
            6062..7719
                /note="assembly_fragment"
            7820..9799
                /note="assembly_fragment"
            9900..11434
                /note="assembly_fragment"
            11535..14382
                /note="assembly_fragment"
            14483..17148
                /note="assembly_fragment"
            17249..19239
                /note="assembly_fragment"
            19340..21102
                /note="assembly_fragment"
            21203..23371
                /note="assembly_fragment"
```

```
misc_feature 23472..25782
    /note="assembly_fragment"
misc_feature 25883..28824
    /note="assembly_fragment"
misc_feature 28925..31619
    /note="assembly_fragment"
misc_feature 31720..34679
    /note="assembly_fragment"
misc_feature 34780..37527
    /note="assembly_fragment"
misc_feature 37628..40146
    /note="assembly_fragment"
misc_feature 40247..43743
    /note="assembly_fragment"
misc_feature 43844..47133
    /note="assembly_fragment"
misc_feature 47234..51023
    /note="assembly_fragment"
misc_feature 51124..54935
    /note="assembly_fragment"
    clone_end:T7
    vector_side:right
misc_feature 55036..59553

Query Match 100.0%; Score 20; DB 2; Length 145833;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTTG 20
    |||||
Db 100045 CTGGACAGGTTAGGCGTTTG 100026

RESULT 9
AC093247 AC093247 80828 bp DNA linear PRI 26-FEB-2002
LOCUS Homo sapiens chromosome 5 clone RP11-138M1, complete sequence.
AC093247 AC093247
VERSION AC093247.3 GI:18921295
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 80828)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            Direct Submission
            Unpublished
            2 (bases 1 to 80828)
            DOE Joint Genome Institute.
            Direct Submission
            Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            3 (bases 1 to 80828)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            Direct Submission
            Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
            On Feb 26, 2002 this sequence version replaced gi:15290461.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www.shgc.stanford.edu
            Quality: Phrap Quality >=40 99.5% of Sequence;
            Estimated Total Number of Errors is 0.4.
            NOTE: This insert is not the entire sequence of the clone (entire
            sequence is 147.2kb). It is clipped at the overlap with AC010237.
            The number of bases overlapped is 38259.

FEATURES
    source
        Location/Qualifiers
            1..80828
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
```

```

source
1. 100300
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-388B4"
/clone_lib="RPC1 human BAC library 11"

ORIGIN
Query Match      87.0%; Score 17.4; DB 9; Length 80828;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTT 19
    |||||
DB 71309 CTGGACAGGTCAGGCGTTT 71327

RESULT 10
AC114970      100300 bp      DNA      linear      HTG 14-MAR-2002
LOCUS      Homo sapiens chromosome 5 clone RP11-388B4, WORKING DRAFT SEQUENCE,
DEFINITION      7 unordered pieces.
AC114970
VERSION      AC114970.1 GI:19424431
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 100300)
AUTHORS      DOE Joint Genome Institute.
TITLE      Sequencing of Human Chromosome 5
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 100300)
AUTHORS      DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL      Submitted (14-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT      -----
Center: JGI
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 555385
Center clone name: RPC1-11_388B4
-----
Summary Statistics
Consensus quality: 95359 bases at least Q40
Consensus quality: 97011 bases at least Q30
Consensus quality: 97825 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Quality coverage: 16.88 in Q20 bases; agarose-fp estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1595: contig of 1595 bp in length
* 1696: gap of unknown length
* 1696: contig of 1499 bp in length
* 3195: gap of unknown length
* 3295: contig of 4404 bp in length
* 7698: gap of unknown length
* 7793: contig of 11003 bp in length
* 18901: gap of unknown length
* 32519: contig of 13618 bp in length
* 32520: gap of unknown length
* 32620: contig of 4976 bp in length
* 37595: gap of unknown length
* 37695: contig of 62605 bp in length.
* 37695: 100300: contig of 62605 bp in length.
Location/Qualifiers

FEATURES
Location/Qualifiers
1595: contig of 1595 bp in length
1696: gap of unknown length
1696: contig of 1499 bp in length
3195: gap of unknown length
3295: contig of 4404 bp in length
7698: gap of unknown length
7793: contig of 11003 bp in length
18901: gap of unknown length
32519: contig of 13618 bp in length
32520: gap of unknown length
32620: contig of 4976 bp in length
37595: gap of unknown length
37695: contig of 62605 bp in length.
37695: 100300: contig of 62605 bp in length.
Location/Qualifiers

ORIGIN
Query Match      87.0%; Score 17.4; DB 2; Length 100300;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCGTTT 19
    |||||
DB 64854 CTGGACAGGTCAGGCGTTT 64872

RESULT 11
AC023817      108742 bp      DNA      linear      HTG 07-MAR-2002
LOCUS      Homo sapiens chromosome 5 clone CTD-2565N12, WORKING DRAFT
DEFINITION      SEQUENCE, 3 ordered pieces.
AC023817
VERSION      AC023817.5 GI:19224766
KEYWORDS      HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 108742)
AUTHORS      DOE Joint Genome Institute.
TITLE      Sequencing of Human Chromosome 5
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 108742)
AUTHORS      DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL      Submitted (18-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT      -----
Center: JGI
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 815265
Center clone name: CTD-El_2565N12
-----
Summary Statistics
Consensus quality: 107872 bases at least Q40
Consensus quality: 108273 bases at least Q30
Consensus quality: 108423 bases at least Q20
Estimated insert size: 137000; agarose-fp estimation
Quality coverage: 8.13 in Q20 bases; agarose-fp estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 31413: contig of 31413 bp in length
* 31414: gap of unknown length

```

\* 31514 101416: contig of 69903 bp in length  
 \* 101417 101516: gap of unknown length  
 \* 101517 108742: contig of 7226 bp in length.

FEATURES  
 source  
 1. 108742  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="CTD-2565N12"  
 /clone\_lib="CalTech human BAC library D"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 108742;  
 Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTT 19  
 |||||  
 Db 70123 CTGGACAGTCAGGCTTT 70141

## RESULT 12

AC139485.1

WPCOMMENT

Sequence split into 4 fragments LOCUS AC139485 Accession AC139485  
 Fragment Name Begin End  
 AC139485\_0 1 110000  
 AC139485\_1 100001 210000  
 AC139485\_2 310000 310000  
 AC139485\_3 300001 368755  
 Continuation (2 of 4) of AC139485 from base 100001 (AC139485 Homo sapiens chromosome 5 c

Query Match 87.0%; Score 17.4; DB 2; Length 110000;  
 Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTT 19

Db 29275 CTGGACAGTCAGGCTTT 29293

## RESULT 13

AC132801/c

LOCUS

AC132801 Homo sapiens chromosome 5 clone CTC-202015, linear PRI 13-NOV-2002

DEFINITION Homo sapiens chromosome 5 clone CTC-202015, complete sequence.

AC132801

AC132801.2 GI:24942889

HTG.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 116702)

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission

Unpublished

2 (bases 1 to 116702)

DOE Joint Genome Institute.

Direct Submission

Submitted (04-SEP-2002) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

3 (bases 1 to 116702)

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission

Submitted (13-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell  
 Drive, Walnut Creek, CA 94598, USA

On Nov 13, 2002 this sequence version replaced gi:22711579.

Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center

www.shgc.stanford.edu

Quality: Phrap Quality &gt;=40 100% of Sequence;

Estimated Total Number of Errors is 0.

## FEATURES

source

1. 116702  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="CTC-202015"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 116702;  
 Best Local Similarity 94.7%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTT 19

|||||

Db 60842 CTGGACAGTCAGGCTTT 60824

## RESULT 14

AC138851/c

LOCUS

AC138851 Homo sapiens chromosome 5 clone RP11-1310P17, linear HTG 21-JAN-2003

DEFINITION Homo sapiens chromosome 5 clone RP11-1310P17, WORKING DRAFT

SEQUENCE, 13 unordered pieces.

AC138851

AC138851.1 GI:27805263

HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 129218)

DOE Joint Genome Institute.

Sequencing of Human Chromosome 5

Unpublished

2 (bases 1 to 129218)

DOE Joint Genome Institute.

Direct Submission

Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

-----

Project Information

Center Project Name: 2723051

Center clone name: RPCI-11\_1310P17

-----

Summary Statistics

Consensus quality: 124109 bases at least Q40

Consensus quality: 125106 bases at least Q30

Consensus quality: 125993 bases at least Q20

Estimated insert size: 175000; agarose-fp estimation

Estimated insert size: 128018; sum-of-contigs estimation

Quality coverage: 8.8 in Q20 bases; agarose-fp estimation

Quality coverage: 12.03 in Q20 bases; sum-of-contigs estimation.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 13 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1550: contig of 1550 bp in length

\* 1551 1650: gap of unknown length

\* 1651 7729: contig of 6079 bp in length

\* 7730 7829: gap of unknown length

\* 7830 9797: contig of 1968 bp in length

\* 9798 9897: gap of unknown length

\* 9898 12559: contig of 2662 bp in length

\* 12560 12659: gap of unknown length

```

* 12650 20867: contig of 8208 bp in length
* 20868 20967: gap of unknown length
* 20968 24297: contig of 3330 bp in length
* 24298 24397: gap of unknown length
* 24398 35099: contig of 10702 bp in length
* 35100 35199: gap of unknown length
* 35200 39384: contig of 4185 bp in length
* 39385 39484: gap of unknown length
* 39485 53302: contig of 19818 bp in length
* 53303 59402: gap of unknown length
* 59403 64594: contig of 5192 bp in length
* 64595 64694: gap of unknown length
* 64695 93174: contig of 28480 bp in length
* 93175 93274: gap of unknown length
* 93275 102519: contig of 9245 bp in length
* 102520 102619: gap of unknown length
* 102620 129218: contig of 26599 bp in length.
FEATURES             Location/Qualifiers
     source           1..129218
     organism="Homo sapiens"
     mol_type="genomic DNA"
     db_xref="taxon:9606"
     chromosome="5"
     clone="RP11-1310P17"
     clone_lib="RPCI human BAC library 11"

```

## ORIGIN

```

Query Match          87.0%; Score 17.4; DB 2; Length 129218;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTT 19
|||||
Db 124866 CTGGACAGGTCAGGGCTTT 124848
|||||

```

## RESULT 15

```

AC138930
LOCUS          133181 bp DNA linear HTG 21-JAN-2003
DEFINITION    Homo sapiens chromosome 5 clone RP11-678014, WORKING DRAFT
SEQUENCE      5 unordered pieces.
AC138930
VERSION       AC138930.1 GI:27805342
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 133181)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 133181)
DOE Joint Genome Institute.
Direct Submission
Submitted (21-JAN-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1555206
Center clone name: RPCI-11_678014
-----
Summary Statistics
Consensus quality: 131745 bases at least Q40
Consensus quality: 131928 bases at least Q30
Consensus quality: 132052 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 132781; sum-of-contigs estimation
Quality coverage: 13.96 in Q20 bases; agarose-fp estimation

```

Quality coverage: 18.4 in Q20 bases; sum-of-contigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

* 1 1094: contig of 1094 bp in length
* 1095 1194: gap of unknown length
* 1195 2421: contig of 1227 bp in length
* 2422 2521: gap of unknown length
* 2522 4100: contig of 1579 bp in length
* 4101 4200: gap of unknown length
* 4201 23545: contig of 19345 bp in length
* 23546 133181: contig of unknown length
* 23646 133181: contig of 109536 bp in length.
FEATURES             Location/Qualifiers
     source           1..133181
     organism="Homo sapiens"
     mol_type="genomic DNA"
     db_xref="taxon:9606"
     chromosome="5"
     clone="RP11-678014"
     clone_lib="RPCI human BAC library 11"

```

## ORIGIN

```

Query Match          87.0%; Score 17.4; DB 2; Length 133181;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTT 19
|||||
Db 121523 CTGGACAGGTCAGGGCTTT 121541
|||||

```

Search completed: February 19, 2004, 23:23:17  
 Job time : 91.562 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:15:48 ; Search time 18.5292 Seconds  
(without alignments)  
4585.415 Million cell updates/sec

Title: US-09-939-853A-140

Perfect score: 20

Sequence: 1 ctggacaggttaggcttgg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_29Jan04.\*

1: geneseq1980s.\*

2: geneseq1990s.\*

3: geneseq2000s.\*

4: geneseq2001as.\*

5: geneseq2001bs.\*

6: geneseq2002s.\*

7: geneseq2003as.\*

8: geneseq2003bs.\*

9: geneseq2003cs.\*

10: geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6	ABK61505 Human NOV
2	20	100.0	444	6	ABQ98669 Human ORF
3	20	100.0	445	5	AAS74747 DNA encod
4	20	100.0	763	6	ABQ98670 Human ORF
5	20	100.0	1183	6	ABK61465 Human CDN
6	20	100.0	2567	6	AAD43980 Human Src
7	16.8	84.0	538	9	ADD22694 Filamento
8	16.8	84.0	589	7	ABZ56309 Aspergill
9	16.8	84.0	809	3	Aaf14531 Aspergill
10	16.8	84.0	2305	7	ADA33227 Human cod
11	16.8	84.0	44325	8	ADA02960 Human DAD
12	16.8	84.0	44325	9	ADB72698 Human DAD
13	16.8	84.0	44325	9	ADC95440 Human Dad
14	16.4	82.0	2446	9	ADE57896 Human gen
15	16.4	82.0	2446	9	ADE57892 Human gen
16	16.4	82.0	2446	9	ADE57900 Human gen
17	16.4	82.0	2446	9	ADE57904 Human gen
18	16.4	82.0	2446	9	ADE83385 Human gen
19	15.8	79.0	403	8	ACH48692 Human leu
20	15.8	79.0	467	8	ACH40016 Human foe
21	15.8	79.0	712	7	ACD95832 Human col
22	15.8	79.0	810	5	ABA16524 Human ner
23	15.8	79.0	1152	2	AAX14434 H. pylori

C 24	15.8	79.0	1255	9	ADE56623 Human gen
C 25	15.8	79.0	1255	9	ADE56627 Human gen
C 26	15.8	79.0	1255	9	ADE56615 Human gen
C 27	15.8	79.0	1255	9	ADE56619 Human gen
C 28	15.8	79.0	1774	2	AAZ42126 Human nor
C 29	15.8	79.0	3153	2	AAV06251 Human pur
C 30	15.8	79.0	3362	2	AAV06278 Human pur
C 31	15.8	79.0	3747	2	AAV06250 Human pur
C 32	15.8	79.0	4049	4	AAH02878 Human she
C 33	15.8	79.0	4170	4	AAO09546 Human pro
C 34	15.8	79.0	4170	10	ADE77056 Human CDN
C 35	15.8	79.0	24053	2	AAI12308 Toxoplasma
C 36	15.8	79.0	32768	2	AAI20515 Polynucle
C 37	15.4	77.0	673	6	ABK63392 Rat seque
C 38	15.4	77.0	673	9	ABK63392 Rat seque
C 39	15.4	77.0	2067	4	ABK63392 Rat seque
C 40	15.4	77.0	2067	6	ABK63392 Rat seque
C 41	15.4	77.0	3685	6	ABK63392 Rat seque
C 42	15.4	77.0	3685	7	ABK63392 Rat seque
C 43	15.4	77.0	4112	5	AAH81792 Human dif
C 44	15.4	77.0	4137	3	AAK76337 Human ORF
C 45	15.4	77.0	6029	4	AAK82712 Human imm

## ALIGNMENTS

### RESULT 1

ABK61505

ID ABK61505 standard; DNA; 20 BP.

XX

AC ABK61505;

XX

DT 18-JUN-2002 (first entry)

XX

DE Human NOV13 RT-PCR primer #1.

XX

KW Human; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis; cell signal processing disorder; metabolic pathway modulation disorder; diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; primer; uterus cancer; immune response; graft-versus-host disease; Exon linking; acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease; hypertension; congenital heart defects; multiple sclerosis; inflammation; Albricht hereditary osteodystrophy; reverse transcriptase PCR.

XX Homo sapiens.

XX

PN WO200216599-A2.

XX

PD 28-FEB-2002.

XX

PF 27-AUG-2001; 2001WO-US026510.

XX

PR 25-AUG-2000; 2000US-0228191P.

PR

PR 08-FEB-2001; 2001US-0267300P.

PR

PR 20-FEB-2001; 2001US-0269961P.

PR

PR 20-MAR-2001; 2001US-0277337P.

XX

(CURA-) CURAGEN CORP.

PA (CORT-) COR THERAPEUTICS INC.

XX

PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;

PI Spytek KA, Szekeres ES, Tomlinson JB, Topper JN, Yang R;

XX

DR WPI; 2002-280937/32.

PT

PT New polypeptides for treating or preventing a disorder associated with them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.

XX

PS Example 2; Page 234; 263pp; English.

XX

CC The invention relates to an isolated polypeptide (NOVX) a mature form of NOVX, a NOVX variant (differing by no more than 15%), the nucleotide

CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,  
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it  
CC and antibody against it, are useful for treating or preventing (e.g. by  
CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
CC atherosclerosis, a disorder related to cell signal processing and  
CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
CC and nucleic acids are also useful for determining the presence of  
CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
CC especially useful in therapeutic or prophylactic applications for  
CC disorders associated with aberrant NOVX expression or activity, e.g.  
CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
CC cancer), immune response, graft-versus-host disease, acquired  
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
CC congenital heart defects, multiple sclerosis, inflammation or Albrit  
CC hereditary osteodystrophy and many other diseases listed in the  
CC specification. The DNA encoding the protein is useful in gene therapy for  
CC treating the conditions. This is also useful in detection assays,  
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
CC for developing a powerful assay system for functional analysis of various  
CC human disorders, as well as in diagnostic applications. The present  
CC sequence is a reverse transcriptase (RT)-PCR primer used to measure  
CC tissue specific expression of mRNA encoding a NOVX protein

XX  
SQ Sequence 20 BP; 3 A; 3 C; 8 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTAGGGCTTTG 20  
Db 1 CTGGACAGGTAGGGCTTTG 20

## RESULT 2

ID ABQ98669/c  
ID ABQ98669 standard; DNA; 444 BP.

AC ABQ98669;

DT 04-NOV-2002 (first entry)

DE Human ORF476 coding sequence.

KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnervary;  
KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
KW cancer; cardiovascular disease; allergy; autoimmune disease;  
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.

OS Homo sapiens.

PN US2002082206-A1.

PD 27-JUN-2002.

PF 30-MAY-2001; 2001US-00867550.

PR 30-MAY-2000; 2000US-0208427P.

PA (LEAC/) LEACH M. D.

PA (MEHR/) MEHRABAN F.

PA (CONL/) CONLEY P. B.

PA (TOPP/) TOPPER J. N.

PA (LAWD/) LAW D.

PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

XX WPI; 2002-626554/67.

DR P-PSDB; ABP64106.

XX New polypeptide designated ORFX are present in human atherogenic cells  
PT and are useful to prevent and treat ORFX-associated disorders including

PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
PT inflammatory disease.

XX Claim 2; SEQ ID NO 951; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their  
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
CC were discovered in human atherogenic cells, in particular in platelets  
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
CC many other tissues as well. Atherogenic cells are cells which have the  
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
CC nucleic acids are useful for treating or preventing a pathological  
CC condition associated with an ORFX-associated disorder, e.g. cancer,  
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
CC coagulation disorders or inflammatory disorders. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from the USPTO web site at  
CC seqdata.uspto.gov/sequence.html?DocID=20020082206

XX SQ Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 444;  
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTAGGGCTTTG 20

Db 60 CTGGACAGGTAGGGCTTTG 41

## RESULT 3

AAS74747

ID AAS74747 standard; cDNA; 445 BP.

XX AAS74747;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #10551.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG10560.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 1; SEQ ID NO 10551; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activities. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAG64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 445 BP; 89 A; 112 C; 143 G; 101 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 5; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20  
 |||||  
 Db 168 CTGGACAGGTTAGGCTTTG 187

RESULT 4  
 ABQ98670/c  
 ID ABQ98670 standard; DNA; 763 BP.  
 XX AC ABQ98670;  
 XX DT 04-NOV-2002 (first entry)  
 XX DE Human ORF477 coding sequence.  
 XX KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vascular;  
 KW Antinflammatory; gene therapy; human; ORFX; atherogenic; platelet;  
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KW cancer; cardiovascular disease; allergy; autoimmune disease;  
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.

XX OS Homo sapiens.  
 XX PN US2002082206-A1.  
 XX PD 27-JUN-2002.  
 XX PF 30-MAY-2001; 2001US-00867550.  
 XX PR 30-MAY-2000; 2000US-0208427P.  
 XX PA (LEAC/) LEACH M D.  
 XX PA (MEHR/) MEHRABAN F.  
 XX PA (CONL/) CONLEY P B.  
 XX PA (TOPP/) TOPPER J N.  
 XX PA (LAWD/) LAW D.  
 XX PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX DR WPI: 2002-626554/67.  
 XX DR P-PSDB; ABP64107.

XX New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.

XX Claim 2; SEQ ID NO 953; 789p; English.  
 XX The present invention relates to novel human ORFX polypeptides and their

CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?DocID=20020082206

XX SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 763;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20  
 |||||  
 Db 189 CTGGACAGGTTAGGCTTTG 170

RESULT 5  
 ABK61465/c  
 ID ABK61465 standard; cDNA; 1183 BP.

XX AC ABK61465;  
 XX DT 18-JUN-2002 (first entry)  
 XX DE Human cDNA encoding protein NOV13.

XX KW Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;  
 KW cell signal processing disorder; metabolic pathway modulation disorder;  
 KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;  
 KW uterus cancer; immune response; graft-versus-host disease;  
 KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;  
 KW hypertension; congenital heart defects; multiple sclerosis; inflammation;  
 KW Albright hereditary osteodystrophy.

XX OS Homo sapiens.

XX PN WO200216599-A2.

XX PD 28-FEB-2002.

XX PF 27-AUG-2001; 2001WO-US026510.

XX PR 25-AUG-2000; 2000US-0228191P.

XX PR 08-FEB-2001; 2001US-0267300P.

XX PR 20-FEB-2001; 2001US-0269961P.

XX PR 20-MAR-2001; 2001US-0277337P.

XX PA (CURA-) CURAGEN CORP.

XX PA (COR-) COR THERAPEUTICS INC.

XX PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;

XX PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;

XX DR WPI: 2002-280937/32.

XX DR P-PSDB; AAU91308.

XX New polypeptides for treating or preventing a disorder associated with

PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.

XX Claim 1; Page 98; 263pp; English.

XX The invention relates to an isolated polypeptide (NOVX) a mature form of

CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide

CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,

15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it

CC and antibody against it, are useful for treating or preventing (e.g. by  
 CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,  
 CC atherosclerosis, a disorder related to cell signal processing and  
 CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide  
 CC and nucleic acids are also useful for determining the presence of  
 CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are  
 CC especially useful in therapeutic or prophylactic applications for  
 CC disorders associated with aberrant NOVX expression or activity, e.g.  
 CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus  
 CC cancer), immune response, graft-versus-host disease, acquired  
 CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,  
 CC congenital heart defects, multiple sclerosis, inflammation or Albrit  
 CC hereditary osteodystrophy and many other diseases listed in the  
 CC specification. The DNA encoding the protein is useful in gene therapy for  
 CC treating the conditions. This is also useful in detection assays,  
 CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or  
 CC for developing a powerful assay system for functional analysis of various  
 CC human disorders, as well as in diagnostic applications. The present  
 CC sequence encodes a NOVX protein  
 XX  
 SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 1183;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20  
 |||||  
 DB 301 CTGGACAGGTTAGGCTTTG 282

RESULT 6  
 AAD43980/c  
 ID AAD43980 standard; cDNA; 2567 BP.

XX AAD43980;

XX 13-DEC-2002 (first entry)

XX Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.

XX Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;  
 KW SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;  
 KW neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;  
 KW psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;  
 KW Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;  
 KW multiple sclerosis; asthma; acute respiratory distress syndrome;  
 KW pulmonary disorder; dermatological; neuroprotective; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH 415..1200  
 CDS /\*tag= a  
 FT /product= "Human SLAP-2"

XX WO200242457-A1.

XX 30-MAY-2002.

XX 20-NOV-2001; 2001WO-US043367.

XX 22-NOV-2000; 2000US-0252545P.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;

XX WPI; 2002-463632/49.

XX P-PSDB; AAE26357.

XX Novel substantially purified human SH2/SH3-domain-containing adapter  
 PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic

PT intervention in immunological and inflammatory disorders and cancer.  
 XX Claim 2; Fig 1; 85pp; English.  
 PS  
 CC The invention relates to a substantially purified human SH2/SH3-domain-  
 CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-  
 CC 2). The invention is useful for treating an immune disorder involving  
 CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is  
 CC useful for screening for antagonists or inhibitors of the interaction of  
 CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or  
 CC preventing diseases or disorders associated with aberrant or uncontrolled  
 CC cellular signal transduction, for determining those cellular signalling  
 CC molecules which associate with hSLAP-2 and which provide critical signals  
 CC for cell activation, and as effectors in methods to affect T- cell  
 CC activation. The invention is useful in screening assays to identify and  
 CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for  
 CC potential use to treat autoimmune diseases which may be caused by  
 CC hyperactivated B cells, as well as to treat diseases which may be caused  
 CC by hyperactivated T cells, in addition to other immune system related  
 CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,  
 CC inflammation disorders, diseases and conditions, rheumatoid arthritis,  
 CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's  
 CC and ulcerative colitis), allergies, particularly those involving  
 CC hyperactivity of B-cells and T- cells, or other immune cells, such as  
 CC mast cells or eosinophils, autoimmune diseases such as systemic lupus  
 CC erythematosus and multiple sclerosis, pulmonary diseases including  
 CC asthma, acute respiratory distress syndrome, and chronic obstructive  
 CC pulmonary disorder, tissue/ organ rejection and cancer. The invention is  
 CC useful in gene therapy. The present sequence is human SLAP-2 cDNA  
 XX

SQ Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 2567;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGCTTTG 20

|||||  
 DB 318 CTGGACAGGTTAGGCTTTG 299

RESULT 7

ADD22694/c

ID ADD22694 standard; DNA; 538 BP.

XX ADD22694;

XX 15-JAN-2004 (first entry)

XX Filamentous fungi Aspergillus genus DNA sequence, SEQ ID No 145.

XX promoter activity; filamentous fungi Aspergillus genus;  
 KW transcription activation; glucose; gene transcription; ds.

OS Aspergillus oryzae.

XX JP2003144171-A.

XX 20-MAY-2003.

XX 16-NOV-2001; 2001JP-00351368.

XX 16-NOV-2001; 2001JP-00351368.

XX (HGET ) HIGETA SHOYU KK.

PA (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.

XX WPI; 2003-818168/77.

XX Novel promoter DNA derived from Aspergillus genus, useful for protein  
 PT production.

PS Claim 1; SEQ ID NO 145; 83pp; Japanese.

XX CC The invention relates to a novel DNA sequence comprising any one of 74  
CC sequences with promoter activity, having a fully defined sequence shown  
CC in the specification or having a sequence which hybridised under  
CC stringent conditions to one of the 74 sequences. The invention further  
CC comprises one of the 74 DNA sequences or their DNA fragments, derived  
CC from filamentous fungi *Aspergillus* genus, being useful for manufacturing  
CC a protein. One of the DNA sequences or its fragment can effectively  
CC activate transcription of a gene encoding a target protein irrespective  
CC of the presence or absence of glucose. The DNA sequences and their  
CC fragments can constantly activate gene transcription. This polynucleotide  
CC sequence represents one of the 74 filamentous fungi *Aspergillus* genus DNA  
CC sequences of the invention.

XX SQ Sequence 538 BP; 128 A; 123 C; 132 G; 155 T; 0 U; 0 Other;  
Query Match 84.0%; Score 16.8; DB 9; Length 538;  
Best Local Similarity 90.0%; Pred. No. 1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20  
|||||  
DB 494 CTGGACAGGTTAGGGCTTTG 475

RESULT 8  
ABZ56309/c  
ID ABZ56309 standard; cDNA; 589 BP.  
XX AC ABZ56309;  
XX DT 28-MAR-2003 (first entry)  
XX DE *Aspergillus oryzae* polynucleotide SEQ ID NO 5422.  
XX KW *Aspergillus oryzae*; fermentation; fungus; industrial; EST;  
XX KW expressed sequence tag; gene; ss.  
XX OS *Aspergillus oryzae*.  
XX XX WO200279476-A1.  
XX PN 10-OCT-2002.  
XX PD 22-MAR-2002; 2002WO-IB000890.  
XX PF 30-MAR-2001; 2001JP-00098371.  
XX PR (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX PA (NARE-) NAT RES INST BREWING.  
XX PA (NORQ) NAT FOOD RES INST MIN AGRIC.  
XX PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;  
XX PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;  
XX WPI; 2003-046817/04.  
XX PT Detection of expression of specific *Aspergillus* genes for monitoring the  
XX fermentation and growth conditions of the fungus, using DNA probes.  
XX PS Claim 1; SEQ ID NO 5422; 48pp + Sequence Listing; Japanese.  
XX CC The invention relates to a polynucleotide having any of 6006 specific  
CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under  
CC specific culture conditions including one or more of eutrophic,  
CC oligotrophic, solid, early germination, alkaline, high temperature, low  
CC temperature or maltose culture or polynucleotides stringently hybridising  
CC to these sequences. The polynucleotides are useful for monitoring the  
CC progress of fermentation and the growth conditions of a fungus,  
CC especially of *Aspergillus oryzae* which is widely used in industrial  
CC fermentation. Also monitoring for fungal contamination. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 589 BP; 143 A; 210 C; 95 G; 141 T; 0 U; 0 Other;  
Query Match 84.0%; Score 16.8; DB 7; Length 589;  
Best Local Similarity 90.0%; Pred. No. 1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20  
|||||  
DB 100 CTGGACAGGTTAGGGCTTTG 81

RESULT 9  
AAF14531/c  
ID AAF14531 standard; cDNA; 809 BP.  
XX AC AAF14531;  
XX DT 13-MAR-2001 (first entry)  
XX DE *Aspergillus oryzae* EST SEQ ID NO:7054.  
XX KW Multiple gene expression; filamentous fungal cell; EST;  
XX KW expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;  
XX KW *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombination;  
XX KW culture condition; environmental stress; spore morphogenesis;  
XX KW metabolic pathway engineering; catabolic pathway engineering; ss.  
XX OS *Aspergillus oryzae*.  
XX XX WO2000056762-A2.  
XX PN 28-SEP-2000.  
XX PD 22-MAR-2000; 2000WO-US007781.  
XX PF 22-MAR-1999; 99US-00273623.  
XX PR (NOVO) NOVO NORDISK BIOTECH INC.  
XX PA (NOVO) NOVO NORDISK AS.  
XX PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
XX WPI; 2000-594572/56.  
XX PT Monitoring differential expression of genes in filamentous fungal cells  
XX PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
XX PT substrate of expressed sequence tags.  
XX PS Claim 88; Page 2863-2864; 3161pp; English.  
XX CC The present invention describes a method for monitoring differential  
XX expression of genes in a first filamentous fungal (FF) cell relative to  
XX expression of the same genes in one or more second filamentous fungal  
XX cells. The method uses fluorescence-labeled nucleic acids isolated from  
XX the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
XX are used in the methods for monitoring differential expression of genes  
XX in a first filamentous fungal (FF) cell relative to expression of the  
XX same genes in one or more second filamentous fungal cells. Monitoring the  
XX global expression of genes from FF cells allows the production potential  
XX of the microorganisms to be improved. New genes may be discovered,  
XX possible functions of unknown open reading frames can be identified and  
XX gene copy number variation and stability can be monitored. The expression  
XX of genes can be used to study how FF cells adapt to changes in culture  
XX conditions, environmental stress, spore morphogenesis, recombination,  
XX metabolic or catabolic pathway engineering. Using ESTs provides several  
XX advantages over genomic or random cDNA clones including elimination of  
XX redundancy as one spot on an array equals one gene or open reading frame,  
XX and organisation of the microarrays based on function of the gene  
XX products to facilitate analysis of the results. AAF07478 to AAF11247  
XX represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents  
XX ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from

CC Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from  
 CC Trichoderma reesei, which are all specifically claimed in the present  
 CC invention  
 XX  
 SQ Sequence 809 BP; 182 A; 265 C; 161 G; 199 T; 0 U; 2 Other;  
 Query Match 84.0%; Score 16.8; DB 3; Length 809;  
 Best Local Similarity 90.0%; Pred. No. 1.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CTGGACAGGTTAGGGCTTTG 20  
 DB 62 CTGGACAGGTTAGGGCTTTG 43

RESULT 10  
 ADA53227/c  
 ID ADA53227 standard; cDNA; 2305 BP.  
 XX  
 AC ADA53227;  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Human coding sequence, SEQ ID 795.  
 XX  
 KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
 KW inflammatory disease; osteoporosis; neurological disease; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1293569-A2.  
 XX  
 PD 19-MAR-2003.  
 XX  
 PF 21-MAR-2002; 2002EP-00006586.  
 XX  
 PR 14-SEP-2001; 2001JP-00328381.  
 PR 24-JAN-2002; 2002US-0350435P.  
 XX  
 PA (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX  
 DR WPI; 2003-395539/38.  
 DR P-P8DB; ADA54866.  
 XX  
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.  
 XX  
 Claim 1; SEQ ID NO 795; 205pp; English.  
 XX  
 The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 XX  
 SQ Sequence 2305 BP; 457 A; 730 C; 621 G; 497 T; 0 U; 0 Other;  
 Query Match 84.0%; Score 16.8; DB 7; Length 2305;  
 Best Local Similarity 90.0%; Pred. No. 1.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CTGGACAGGTTAGGGCTTTG 20  
 DB 2118 CTGGACAGGTTAGGGCTTTG 2099

RESULT 11  
 ADA02960/c  
 ID ADA02960 standard; DNA; 44325 BP.  
 XX  
 AC ADA02960;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE Human DAD1 carcinoma associated gene, SEQ ID NO:1478.  
 XX  
 KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;  
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
 KW gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003057146-A2.  
 XX  
 PD 17-JUL-2003.  
 XX  
 PF 26-DEC-2002; 2002WO-US041414.  
 XX  
 PR 26-DEC-2001; 2001US-00035832.  
 XX  
 PA (SAGR-) SAGRES DISCOVERY.  
 XX  
 PI Morris DW;  
 XX  
 DR WPI; 2003-597068/55.  
 XX  
 PT New recombinant nucleic acid encoding carcinoma associated protein,  
 PT useful for preparing compositions for treating carcinomas.  
 XX  
 Claim 1; SEQ ID NO 1478; 245pp; English.  
 XX  
 The invention relates to recombinant carcinoma associated (CA) nucleic  
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to  
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The  
 CC invention also encompasses expression vectors and host cells comprising a  
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically  
 CC binds to the protein, and a biochip comprising CA nucleic acid or  
 CC fragments thereof. The sequences of the invention were identified using  
 CC oncogenic retroviruses, which insert into the genome of the host organism  
 CC at random. Many of these do not carry transduced host oncogenes or  
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a  
 CC direct consequence of the effects of proviral integration into host  
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
 CC leukaemia) or a propensity to carcinoma by determination of the sequence  
 CC of a CA gene, or by determination of CA gene expression in particular  
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
 CC therapeutic agents and in screening and evaluating drug candidates. The  
 CC present sequence represents a specifically claimed human CA nucleic acid  
 CC sequence of the invention. Note: The complete sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 44325 BP; 11218 A; 9675 C; 9902 G; 13530 T; 0 U; 0 Other;  
 Query Match 84.0%; Score 16.8; DB 8; Length 44325;  
 Best Local Similarity 90.0%; Pred. No. 1.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CTGGACAGGTTAGGGCTTTG 20  
 DB 38075 CTGGACAGGTTAGGGCTTTG 38056

RESULT 12  
 ADB72698/c  
 ID ADB72698 standard; DNA; 44325 BP.  
 XX

```

AC ADB72698;
XX
XX 04-DEC-2003 (first entry)
XX
XX Human DAD1 gene.
XX
XX human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
XX Homo sapiens.
XX
XX WO2003008583-A2.
XX
XX 30-JAN-2003.
XX
XX 26-DEC-2001; 2001WO-US051291.
XX
XX 02-MAR-2001; 2001US-00798596.
XX
XX 23-OCT-2001; 2001US-00004113.
XX
XX 08-NOV-2001; 2001US-00052482.
XX
XX 30-NOV-2001; 2001US-00997722.
XX
XX 20-DEC-2001; 2001US-00034650.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW, Engelhard EK;
XX
XX WPI; 2003-239337/23.
XX
XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
XX Claim 1; SEQ ID NO 526; 2304pp; English.
XX
XX The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a human gene of the invention.
XX
XX Sequence 44325 BP; 11218 A; 9675 C; 9902 G; 13530 T; 0 U; 0 Other;
SQ
Query Match 84.0%; Score 16.8; DB 9; Length 44325;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20
Db 38075 CTGGACAGGATAGGGCTGTG 38056

RESULT 13
ADCS440/c
ID ADC85440 standard; DNA; 44325 BP.
XX
AC ADC85440;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human Dad1 genomic sequence.
XX
XX Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.
XX
XX Homo sapiens.
XX
XX WO2003045230-A2.
XX
XX 05-JUN-2003.
XX
XX 02-DEC-2002; 2002WO-US038582.
XX

```

```

XX
XX 30-NOV-2001; 2001US-00997722.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW, Engelhard EK;
XX
XX WPI; 2003-513603/48.
XX
XX New recombinant nucleic acid comprising a nucleotide sequence of any of
PT the carcinoma-associated (CA) genes, useful for screening for drug
PT candidates for diagnosing or treating carcinomas.
XX
XX Claim 1; SEQ ID NO 226; 983pp; English.
XX
XX The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carcinoma-
CC associated (CA) genes from the 50 tables given in the specification. The
CC CA proteins are secreted, transmembrane or intracellular proteins. The
CC recombinant nucleic acids are useful for screening for drug candidates
CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
XX ADC85514 represent CA genes of the invention.
XX
XX Sequence 44325 BP; 11218 A; 9675 C; 9902 G; 13530 T; 0 U; 0 Other;
SQ
Query Match 84.0%; Score 16.8; DB 9; Length 44325;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20
Db 38075 CTGGACAGGATAGGGCTGTG 38056

RESULT 14
ADES7896
ID ADE57896 standard; DNA; 2446 BP.
XX
AC ADE57896;
XX
DT 29-JAN-2004 (first entry)
XX
XX Human gene U72649, SEQ ID NO 3762.
XX
XX Human; ds; gene; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHC) GEN HOSPITAL CORP.
XX
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX
XX GENBANK; U72649.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX

```

CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human DNA (shown in Table 2 of the  
 CC specification) which encodes one of the polypeptides of the invention  
 CC which is differentially expressed during pain. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 2446 BP; 567 A; 637 C; 596 G; 646 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 9; Length 2446;  
 Best Local Similarity 94.4%; Pred. NO. 1.9e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 GGACAGGTTAGGCTTTG 20  
 |||||  
 DB 1850 GGACAGGTTAGGCTTTG 1867

RESULT 15

AD5E7892

ID AD5E7892 standard; DNA; 2446 BP.

AC AD5E7892;

DT 29-JAN-2004 (first entry)

DE Human gene U72649, SEQ ID NO 3758.

KW Human; ds; gene; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.

OS Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; U72649.

XX

PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human DNA (shown in Table 2 of the  
 CC specification) which encodes one of the polypeptides of the invention  
 CC which is differentially expressed during pain. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 2446 BP; 567 A; 637 C; 596 G; 646 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 9; Length 2446;  
 Best Local Similarity 94.4%; Pred. NO. 1.9e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 GGACAGGTTAGGCTTTG 20

DB 1850 GGACAGGTTAGGCTTTG 1867

Search completed: February 19, 2004, 21:51:42  
 Job time : 21.5292 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:54 ; Search time 132.422 Seconds  
(without alignments)  
4510.152 Million cell updates/sec

Title: US-09-939-853A-140  
Perfect score: 20  
Sequence: 1 ctggacaggttagggcttg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST.\*
- 1: em\_estba.\*
  - 2: em\_esthum.\*
  - 3: em\_estin.\*
  - 4: em\_estnu.\*
  - 5: em\_estov.\*
  - 6: em\_estpl.\*
  - 7: em\_estro.\*
  - 8: em\_esti.\*
  - 9: gb\_esti.\*
  - 10: gb\_est2.\*
  - 11: gb\_hic.\*
  - 12: gb\_est3.\*
  - 13: gb\_est4.\*
  - 14: gb\_est5.\*
  - 15: em\_estfun.\*
  - 16: em\_estom.\*
  - 17: em\_gss\_hum.\*
  - 18: em\_gss\_inv.\*
  - 19: em\_gss\_pln.\*
  - 20: em\_gss\_vrt.\*
  - 21: em\_gss\_fun.\*
  - 22: em\_gss\_mam.\*
  - 23: em\_gss\_mus.\*
  - 24: em\_gss\_pro.\*
  - 25: em\_gss\_rod.\*
  - 26: em\_gss\_pbg.\*
  - 27: em\_gss\_vrl.\*
  - 28: gb\_gss1.\*
  - 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	616	13	BX383606
C 2	20	100.0	778	12	BG178487
C 3	20	100.0	878	12	BQ053486
C 4	20	100.0	986	12	BQ054265

C 5	20	100.0	1020	12	BQ054281
C 6	20	100.0	1201	9	AL541041
C 7	18.4	92.0	901	13	BU327205
C 8	18.4	92.0	962	14	CA987486
C 9	17.4	87.0	332	9	AA719167
C 10	17.4	87.0	357	9	AA759254
C 11	17.4	87.0	413	13	BX089099
C 12	17.4	87.0	740	29	AG048198
C 13	16.8	84.0	177	29	CG591538
C 14	16.8	84.0	212	14	CB487227
C 15	16.8	84.0	259	10	BB306076
C 16	16.8	84.0	316	14	CB486524
C 17	16.8	84.0	380	29	AG214612
C 18	16.8	84.0	415	10	BF665847
C 19	16.8	84.0	470	10	BF448079
C 20	16.8	84.0	498	28	BH258504
C 21	16.8	84.0	545	12	BI529969
C 22	16.8	84.0	559	13	BU648008
C 23	16.8	84.0	624	29	CE663064
C 24	16.8	84.0	628	14	CB445363
C 25	16.8	84.0	645	12	BM003795
C 26	16.8	84.0	695	29	CE509914
C 27	16.8	84.0	721	14	CF206129
C 28	16.8	84.0	727	10	BF215517
C 29	16.8	84.0	749	29	AG053044
C 30	16.8	84.0	757	28	BH012585
C 31	16.8	84.0	907	10	BF681379
C 32	16.8	84.0	942	10	BF310891
C 33	16.8	84.0	989	10	BE729761
C 34	16.8	84.0	990	10	BF699088
C 35	16.4	82.0	206	29	CG621536
C 36	16.4	82.0	223	9	AU254349
C 37	16.4	82.0	283	10	AW799548
C 38	16.4	82.0	297	29	CE360042
C 39	16.4	82.0	363	14	CD995897
C 40	16.4	82.0	368	28	AQ423959
C 41	16.4	82.0	378	9	AV208185
C 42	16.4	82.0	423	10	AW558403
C 43	16.4	82.0	541	28	AZ700347
C 44	16.4	82.0	577	29	CNS07H2B
C 45	16.4	82.0	578	28	AZ850663

## ALIGNMENTS

RESULT 1  
BX383606/c 616 bp mRNA linear EST 08-MAY-2003

LOCUS BX383606 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED

DEFINITION Homo sapiens cDNA clone CS00J013YK10 5-PRIME, mRNA sequence.

ACCESSION BX383606

VERSION BX383606.1 GI:30457152

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 616)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9825.r For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS00J013BF05OPI&cluster=9825.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ invitrogen Corporation 1500

Paraday Avenue Genoscope sequence ID : CS0DJ013BF05Q21.

## FEATURES

source

1. .616  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DJ013VK10"  
/cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/cell\_line="JURKAT"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 616;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGTTAGGGCTTTG 20

Db 350 CTGCACAGGTTAGGGCTTTG 331

## RESULT 2

BG178487/c

LOCUS

DEFINITION BG178487 778 bp mRNA linear EST 06-FEB-2001  
602328305F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4429896 5';  
mRNA sequence.

ACCESSION BG178487

VERSION BG178487.1 GI:12685190

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 778)

TITLE NIH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAW10182 row: 1 column: 01

High quality sequence stop: 657.

Location/Qualifiers

1. .778

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4429896"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 91"

/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.4 Kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 778;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGTTAGGGCTTTG 20

Db 265 CTGCACAGGTTAGGGCTTTG 246

## RESULT 3

BQ053486/c

LOCUS

DEFINITION BQ053486 878 bp mRNA linear EST 29-MAR-2002

AGENCOURT\_6820248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5935253

5'; mRNA sequence.

ACCESSION BQ053486

VERSION BQ053486.1 GI:19812826

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 878)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHC2122 row: 1 column: 06

High quality sequence stop: 394.

Location/Qualifiers

1. .878

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5935253"

/tissue\_type="natural killer cells, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 106"

/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GCACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 878;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGTTAGGGCTTTG 20

Db 176 CTGCACAGGTTAGGGCTTTG 157

## RESULT 4

BQ054265/c

LOCUS

DEFINITION BQ054265 986 bp mRNA linear EST 29-MAR-2002

AGENCOURT\_6830248 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936339

5'; mRNA sequence.

ACCESSION BQ054265

VERSION BQ054265.1 GI:19813605

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 986)

**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LICM2125 row: i column: 12  
 High quality sequence stop: 515.

**FEATURES**  
 source  
 1..985  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5936339"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 986;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTTAGGGCTTTG 20  
 Db 182 CTGGACAGGTTAGGGCTTTG 163

**RESULT 5**  
**LOCUS** BQ054281.1 1020 bp mRNA linear EST 29-MAR-2002  
**DEFINITION** AGENCOURT 6830234 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5936362  
 5', mRNA sequence.  
**ACCESSION** BQ054281  
**VERSION** BQ054281.1 GI:19813621  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE**  
**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LICM2125 row: j column: 11  
 High quality sequence stop: 556.

## FEATURES

source  
 1..1020  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"

/db\_xref="taxon:9606"  
 /clone="IMAGE:5936362"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_106"  
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 1020;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTTAGGGCTTTG 20  
 Db 115 CTGGACAGGTTAGGGCTTTG 96

## RESULT 6

**AL541041/c**  
**LOCUS** AL541041 Homo sapiens 1201 bp mRNA linear EST 12-MAY-2003  
**DEFINITION** 5-PRIME, mRNA sequence.  
**ACCESSION** AL541041  
**VERSION** AL541041.2 GI:30544829  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 1201)  
**AUTHORS** Li W.B., Gruber, C., Jessee, J. and Palayes, D.  
**TITLE** Full-length cDNA libraries and normalization  
**JOURNAL** Unpublished (2001)  
**COMMENT** On Feb 15, 2001 this sequence version replaced gi:12871733.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [secre@genoscope.cns.fr](mailto:secre@genoscope.cns.fr), Web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 9825.r For  
 more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODE005AF12QP&cluster=9825.r>. Contact :  
 Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CSODE005AF12QF1.

## FEATURES

source  
 1..1201  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODE005YK23"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1201;  
 Best Local Similarity 100.0%; Pred. No. 98;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTTAGGGCTTTG 20

```

Db          362 CTGGACAGGTTAGGGCTTTG 343
|||||
RESULT 7
BU327205
LOCUS
DEFINITION
603490679F1 CSEQCHN63 Gallus gallus cDNA clone CHST392m20 5', mRNA
sequence.
ACCESSION
BU327205
VERSION
BU327205.1 GI:25835206
KEYWORDS
EST.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (bases 1 to 901)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE
22335534
PubMed
1244532
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..901
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHST392m20"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN63"
/note="Organ: heads; Vector: pBluescript II KS(+); Site: 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methyated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., FNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Query Match 92.0%; Score 18.4; DB 13; Length 901;
Best Local Similarity 95.0%; Pred. No. Se+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20
|||||
Db 659 CTGGACGGTTAGGGCTTTG 678
|||||

RESULT 8
CA987486
LOCUS
DEFINITION
AGENCOURT 11284551 NICHD XGC Emb1 Xenopus laevis cDNA clone
IMAGE:6864145 5', mRNA sequence.
ACCESSION
CA987486

```

```

VERSION
CA987486.1 GI:27520161
KEYWORDS
EST.
SOURCE
Xenopus laevis (African clawed frog)
ORGANISM
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 962)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabps@mail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14484 row: g column: 24
High quality sequence start: 228
High quality sequence stop: 380.
Location/Qualifiers
1..962
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6864145"
/tissue_type="embryo (stage 10)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb1"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.55 kb. Constructed by Life Technologies. Note: This
is a Xenopus Gene Collection (XGC) library."

ORIGIN
Query Match 92.0%; Score 18.4; DB 14; Length 962;
Best Local Similarity 95.0%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTTG 20
|||||
Db 840 CTGGCAGGTTAGGGCTTTG 821
|||||

RESULT 9
AA719167/c
LOCUS
DEFINITION
ah47e08.s1 Soares_testis_NHT Homo sapiens cDNA clone 1292678 3',
mRNA sequence.
ACCESSION
AA719167
VERSION
AA719167.1 GI:2732266
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 332)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabps@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

```

```

www-bio.llnl.gov/bbrp/image/image.html
Insert length: 472 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 316.
Location/Qualifiers
1..332
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="1292678"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5].
TGTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization to Cots5, and was
constructed by Bento Soares and M. Fatima Bonaudo. "

ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 332;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTAGGGCTTT 19
Db 207 CTGGACAGGTAGGGCTTT 189

RESULT 10
AA759254/c
LOCUS
DEFINITION
ah4942.s1 Soares_testis_NHT Homo sapiens cDNA clone 1322374 3',
mRNA sequence.
ACCESSION
AA759254
VERSION
AA759254.1 GI:2807117
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 357)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaudo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 466 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.
Location/Qualifiers
1..357
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="1322374"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5].
TGTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization to Cots5, and was
constructed by Bento Soares and M. Fatima Bonaudo. "

ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 357;
Best Local Similarity 94.7%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTAGGGCTTT 19
Db 206 CTGGACAGGTAGGGCTTT 188

RESULT 11
BX089099
LOCUS
DEFINITION
BX089099 Soares testis_NHT Homo sapiens cDNA clone IMAGp998P23333
; IMAGE:1322374, mRNA Sequence.
ACCESSION
BX089099
VERSION
BX089099.1 GI:27823687
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 413)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radcliff, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGp998P23333.
RZPLIB; I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/cloneCards/cgi-
bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubneweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.
Location/Qualifiers
1..413
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998P23333 ; IMAGE:1322374"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis_NHT"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5].
TGTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization to Cots5, and was
constructed by Bento Soares and M. Fatima Bonaudo. "

```

```

ORIGIN
constructed by Bento Soares and M. Fatima Bonaldó. "

Query Match      87.0%; Score 17.4; DB 13; Length 413;
Best Local Similarity 94.7%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTGGACAGGTTAGGGCTTT 19
Db 214 CTGGACAGGTCAGGGCTTT 232

RESULT 12
AG048198/c
LOCUS      740 bp      DNA      linear      GSS 02-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-027P05.R, genomic survey sequence.
ACCESSION  AG048198
VERSION     AG048198.1 GI:16585090
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE  1 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
            BAC end sequences of Library PTB
            Unpublished
            2 (bases 1 to 740)
            Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
            Direct Submission
            Submitted (02-AUG-2001) Asso Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-cho,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimbesgsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the Rad process and may have higher chance of
            clone tracking errors.
            PRIMERS
            Sequencing: M13rev
            LIBRARY
            Vector      : pKS145
            R.Site 1   : SacI
            R.Site 2   : SacI
            Location/Qualifiers
            1..740
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-027P05.R"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      87.0%; Score 17.4; DB 29; Length 740;
Best Local Similarity 94.7%; Pred. No. 1.4e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGGACAGGTTAGGGCTTT 20
Db 583 TGGACAGGTTAGGGCTTT 565

RESULT 13
CG591538/c
LOCUS      177 bp      DNA      linear      GSS 02-OCT-2003
DEFINITION OST246195 Mus musculus 129Sv/Ev Mus musculus genomic clone
ACCESSION  CG591538
VERSION     CG591538.1 GI:37399554
KEYWORDS   GSS.
SOURCE     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 177)
            Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggett,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W., Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.F.
            Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
            Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
            Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6576):608-11)
            Class: Gene Trap.
            Location/Qualifiers
            1..177
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /strain="129Sv/Ev"
            /db_xref="taxon:10090"
            /clone="OST246195"
            /cell_type="embryonic stem cell"
            /clone_lib="Mus musculus 129Sv/Ev"

ORIGIN
Query Match      84.0%; Score 16.8; DB 29; Length 177;
Best Local Similarity 90.0%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTGCACAGGTTAGGGCTTTG 20
Db 44 CTGCACAGGATAGGGCTTAG 25

RESULT 14
CB487227/c
LOCUS      212 bp      mRNA      linear      EST 01-APR-2003
DEFINITION ONYKt001012 Oncorhynchus mykiss reproductive Oncorhynchus mykiss
            cDNA, mRNA sequence.
ACCESSION  CB487227
VERSION     CB487227.1 GI:29298453
KEYWORDS   EST.
SOURCE     Oncorhynchus mykiss (rainbow trout)
ORGANISM   Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE  1 (bases 1 to 212)
            GRASP Consortium, Davidson,W.S., Koop,B.F. and
            http://web.uvic.ca/cbr/grasp.
            A survey of Salmo salar transcripts from high complexity cDNA
            libraries
            Unpublished (2002)
            Contact: Koop BF
            Centre for Biomedical Research
            University of Victoria
            PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
            Tel: 250 472 4067
            Fax: 250 472 4075
            Email: Bkoop@uvic.ca
            Centre for Biomedical Research, University of Victoria cDNA
            preparation and sequencing: Roberto Alberto, Marianne
            Beetz-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, BF Koop.

```